

Métodos Quantitativos Aplicados

A Casos Reais



Métodos Quantitativos Aplicados

A Casos Reais

Marco Aurélio Carino Bouzada

© 2013, Elsevier Editora Ltda.

Todos os direitos reservados e protegidos pela Lei 9.610 de 19/02/98. Nenhuma parte deste livro, sem autorização prévia por escrito da editora, poderá ser reproduzida ou transmitida sejam quais forem os meios empregados: eletrônicos, mecânicos, fotográficos, gravação ou quaisquer outros.

Copidesque: Edna da Silva Cavalcanti

Editoração Eletrônica: Thomson Digital

Revisão Gráfica: Lara Alves

Elsevier Editora Ltda.

Conhecimento sem Fronteiras

Rua Sete de Setembro, 111 – 16º andar
20050-006 – Centro – Rio de Janeiro – RJ – Brasil

Rua Quintana, 753 – 8º andar
04569-011 – Brooklin – São Paulo – SP

Serviço de Atendimento ao Cliente
0800-0265340 sac@elsevier.com.br

ISBN: 978-85-352-6808-9

ISBN (versão eletrônica): 978-85-352-7221-5

Nota: Muito zelo e técnica foram empregados na edição desta obra. No entanto, podem ocorrer erros de digitação, impressão ou dúvida conceitual. Em qualquer das hipóteses, solicitamos a comunicação ao nosso Serviço de Atendimento ao Cliente, para que possamos esclarecer ou encaminhar a questão. Nem a editora nem o autor assumem qualquer responsabilidade por eventuais danos ou perdas a pessoas ou bens, originados do uso desta publicação.

**CIP-BRASIL. CATALOGAÇÃO NA PUBLICAÇÃO
SINDICATO NACIONAL DOS EDITORES DE LIVROS, RJ**

M552

Métodos quantitativos aplicados a casos reais / Corso Jansen ... [et al.] ; Marco Aurélio Bouzada (coordenador). - 1. ed. - Rio de Janeiro : Elsevier, 2013.
256 p. : il. ; 23 cm.

ISBN 978-85-352-6808-9

1. Processo decisório. 2. Administração - Metodologia. 3. Pesquisa quantitativa. I. Jansen, Corso. II. Bouzada, Marco Aurélio.

13-00565

CDD: 658.403

CDU: 005.22

29/04/2013 29/04/2013

Dedicatória

Dedico este trabalho aos meus pais, Héctor e Lenise, sem os quais minha moral e educação certamente teriam seguido um caminho que não teria me trazido até aqui, à minha esposa Caroline e ao meu filho Gabriel.

Marco Aurélio Carino Bouzada

Atualmente, as empresas estão muito preocupadas com a redução de custos. Tornar suas operações mais eficientes é um caminho potencialmente capaz de atingir esse objetivo. Para que a eficiência operacional seja alcançada, no entanto, muitas decisões precisam ser tomadas pelo nível gerencial da empresa. Algumas dessas decisões são críticas e suas consequências podem ser formidáveis ou desastrosas no que diz respeito à eficiência da operação.

A busca pelas decisões ótimas nem sempre é um processo simples, mas muitas vezes é encarado como tal, carecendo, em boa parte do emprego de um cientificismo maior. O processo de tomada de decisão é, frequentemente, apoiado apenas pela experiência profissional e intuição do decisor. Desta forma, é de suma importância pesquisar, analisar, criticar, desenvolver e aplicar ao cotidiano operacional das empresas métodos quantitativos de previsão e de apoio à busca pelas decisões ótimas, que objetivam complementar (e não substituir!) a experiência profissional e a intuição do tomador de decisão.

Em outras palavras, é vital explorar ferramentas capazes de embasar as decisões operacionais a serem tomadas pelo administrador, assim como justificá-las e comunicá-las para toda a empresa. Nesse sentido, o avanço das tecnologias computacionais pode facilitar ainda mais a aplicação de métodos quantitativos de previsão e de apoio à tomada de decisão.

Pois a proposta deste livro é justamente explorar essas ferramentas. Para tal, a obra é dividida em duas partes: métodos de apoio à tomada de decisão e métodos de previsão. Em ambas, cada capítulo apresenta um método diferente, com uma revisão conceitual e o passo a passo de sua aplicação a algum caso prático, além da apresentação e discussão de resultados, com as respectivas planilhas de implementação.

Na Parte I, a Programação Linear, as Árvores de Decisão e a Simulação de Monte Carlo são apresentadas teoricamente e aplicadas a problemas reais na área de Administração. O último capítulo dessa parte realiza uma análise comparativa destas metodologias em relação a outro problema real.

A Parte II consiste em dois capítulos que aplicam a Regressão Múltipla (com variáveis *dummy*) a problemas reais de previsão na área de negócios, e um tutorial (com texto e planilha) sobre um complexo método de séries temporais, a Decomposição Clássica.

Este trabalho é destinado a alunos, professores e pesquisadores dos cursos de graduação, pós-graduação e politécnico em Administração, Engenharia de Produção e Logística, inserindo-se perfeitamente como livro-texto para disciplinas como Métodos Quantitativos, Métodos Quantitativos Aplicados e Métodos de Apoio à Decisão. Pode ser também utilizado como bibliografia complementar para disciplinas como Previsão de Vendas, Análise de Decisões, Pesquisa Operacional e Otimização em Sistemas Logísticos.

Os profissionais de mercado também podem usufruir dessa obra para consultas visando aplicações próprias. Analistas de Logística, Inteligência de Marketing e de Planejamento podem encontrar grande utilidade na leitura desta obra para a realização de suas atividades de transportes, previsão de vendas e dimensionamento de equipes.

Desta forma, este trabalho procura preencher uma lacuna existente na interface entre a academia e o mundo dos negócios, apresentando, além do arcabouço teórico, pesquisas acadêmicas (e seus resultados!) recentes na área, totalmente aplicadas e em linha com as metodologias abordadas.

Marco Aurélio Carino Bouzada

Organizador

Fevereiro de 2013

Métodos de apoio à tomada de decisão

Os métodos de apoio à tomada de decisão podem ajudar as empresas a, por exemplo:

- decidir que tipo de tecnologia utilizar para a obtenção de um novo produto, considerando os riscos envolvidos e as consequências potenciais de cada alternativa;
- designar elementos de uma equipe para a realização de diversas tarefas de forma otimizada;
- optar pela modalidade de seguro mais interessante em face dos riscos inerentes ao negócio e aos custos e das recompensas referentes a cada alternativa considerada;
- alocar recursos em diferentes opções de mídia de forma eficiente;
- dimensionar corretamente a produção de determinado item, tendo em mente a flutuação natural da demanda e eventuais custos de sobra e de falta incorridos;
- decidir onde instalar uma antena de transmissão, levando em conta o que pode acontecer se cada um dos possíveis locais for escolhido.

Obviamente, tais decisões podem ser tomadas sem o auxílio de um método de apoio e, lamentavelmente, o são com alguma frequência. Quando isso acontece, no entanto, muitas vezes as consequências de tais decisões tomadas de forma intuitiva ou baseada *apenas* na experiência profissional do gestor podem ser desastrosas:

- a tecnologia escolhida pode não ser a mais indicada para o caso, resultando em custos associados muito maiores que os ganhos;
- funcionários podem ser alocados a tarefas de forma que elas sejam realizadas em um período muito maior do que o necessário;
- o seguro escolhido, por ser o mais barato, pode não cobrir satisfatoriamente os riscos da operação de uma empresa;

- as mídias escolhidas para a alocação de recursos podem alcançar uma quantidade insatisfatória de potenciais consumidores;
- o mal dimensionamento da produção pode fazer com que um item fique “encalhado” no estoque;
- o local escolhido para a instalação de uma antena pode ser, tecnicamente, o mais interessante, mas proibitivamente caro.

Justamente por isso, acaba se tornando um gestor diferenciado aquele que procura minimizar a ocorrência de consequências como as descritas anteriormente, com a utilização de métodos quantitativos para subsidiar a tomada de decisão empresarial. Tais métodos não têm como objetivo substituir o *feeling* e o bom senso do tomador de decisão; apenas complementá-los, dando *apoio* ao processo.

Os métodos quantitativos de apoio à tomada de decisão (que são o objeto desta Parte I) podem ser divididos em dois tipos: os determinísticos e os não determinísticos (estes últimos podendo ser subdivididos em probabilísticos e não probabilísticos).

Os métodos determinísticos só podem ser usados satisfatoriamente nas chamadas decisões em condições de certeza (quando não há dúvida a respeito dos valores das variáveis relevantes). O representante clássico é a Programação Linear, cujo caso particular (Programação Inteira) é explorado no Capítulo 1, que apresenta uma tentativa de solucionar um problema característico de logística da Marinha do Brasil: a alocação, o transporte e a distribuição de gêneros frigorificados para as Organizações Militares.

Já os métodos não determinísticos não probabilísticos são adequados para as decisões em condições de incerteza, que ocorrem quando não se tem praticamente nenhuma informação a respeito de algumas variáveis relevantes, a não ser os seus possíveis valores de ocorrência (mas sem as respectivas probabilidades). As alternativas mais utilizadas em tais situações são os critérios Maximax, Maximin e Minimax.

Por fim, os métodos probabilísticos podem ser aplicados em situações consideradas no meio do caminho entre as condições de certeza e as condições de incerteza: as decisões em condições de risco – quando não se sabe ao certo os valores de algumas variáveis relevantes, mas são conhecidas as suas distribuições de probabilidades. Em tais situações, dois dos métodos mais utilizados são a Árvore de Decisão e a Simulação.

Quando há diferentes momentos em que é necessária uma tomada de decisão, com ocorrência de eventos probabilísticos entre eles, a Árvore de Decisão é mais recomendada. O Capítulo 2 usa esta metodologia para avaliar a escolha de um fornecedor sob condições de risco.

Se, no entanto, a situação compreende muitas complexidades a serem modeladas e/ou não se está preocupado apenas com a decisão ótima, mas também com o desempenho do sistema como um todo em face de cada uma das alternativas de decisão, então a Simulação é mais adequada.

Tal ferramenta é utilizada no Capítulo 3 para verificar se a intermodalidade de transporte é uma alternativa viável no processo de escoamento de arroz.

O Capítulo 4, o último da Parte I, utiliza como pano de fundo um problema real de determinação do mix de projetos de consultoria para comparar a adequação de todas estas metodologias citadas: Programação Linear (de fato, o seu caso particular, a Programação Binária), critérios Maximax, Maximin e Minimax, Árvore de Decisão e Simulação.

A aplicação da Programação Inteira na solução logística do transporte de carga na Marinha do Brasil

Ricardo França Santos

MADE/UNESA

Eugênio Corrêa de Souza Junior

MADE/UNESA

Marco Aurélio Carino Bouzada

MADE/UNESA

Objetivos de aprendizagem do capítulo

- Saber modelar um problema real em termos determinísticos.
- Entender o conceito de variáveis de decisão, função-objetivo e restrições em um modelo de Programação Linear.
- Aprender a usar o Solver do Excel para otimizar modelos de Programação Linear.
- Entender as peculiaridades da Programação Inteira e da Programação Binária.
- Conhecer algumas das limitações do Solver do Excel na busca pela solução ótima.

INTRODUÇÃO

A Marinha do Brasil realiza suas atividades de abastecimento por meio de um sistema constituído por órgãos de distribuição, processos e recursos de qualquer natureza, interligados e interdependentes, estruturado com a finalidade de prover gêneros necessários à manutenção das Organizações Militares (OMs). Para isso conta com depósitos primários, depósitos secundários e depósitos regionais.

Dentre os depósitos regionais existe o Depósito de Subsistência da Marinha no Rio de Janeiro (DepSubMRJ), criado em 1955, cuja missão é executar as tarefas de receber, periciar, estocar, controlar e fornecer gêneros alimentícios, a fim de contribuir para a prontidão operativa dos Meios Navais, Aeronavais, Fuzileiros Navais e demais Organizações Militares da Marinha.

Em relação à função logística de transporte, diversos requisitos devem ser atendidos pelo DepSubMRJ, principalmente em relação à temperatura da carga e do veículo, assim como em relação a outras características deste último, como estado de conservação, de higiene e de segurança.

O DepSubMRJ dispõe de um número limitado de veículos de transporte de cargas que atendem a esses requisitos, que nem sempre estão disponíveis devido a problemas mecânicos ou ausência de motoristas.

A frota de veículos é constituída de caminhões leves e ágeis que carregam cargas menores, e de caminhões pesados para cargas maiores e percursos mais longos. A distribuição da frota é realizada de modo empírico, baseada na experiência e competência dos profissionais, a partir das quais é produzida uma programação semanal, sustentada em fatores como demanda, localização do destinatário, condições de acesso, distância, disponibilidade do motorista e consumo de combustível, entre os principais fatores. Por outro lado, existe uma demanda, por parte das OMs, no que tange aos gêneros frigorificados, que necessitam ser atendidos por essa frota. A programação semanal atende tanto às OMs na Cidade do Rio de Janeiro como às das cidades próximas. As entregas são realizadas de segunda-feira a sexta-feira, inclusive feriados. Na **Tabela 1.1** estão listados os custos fixos e variáveis da operação.

Na **Tabela 1.2** estão registrados os valores mínimo e máximo de entrega diária.

Na **Tabela 1.3** podem ser observadas a distância e a demanda semanal de cada OM.

O transporte representa um significativo custo logístico por parte das empresas, absorvendo de um a dois terços do total. Quando realizado por frota e equipamentos próprios, pode propiciar melhor desempenho operacional, maior disponibilidade e

TABELA 1.1 Custos fixos e variáveis

Custos fixos (em R\$/mês)	Custos variáveis (em R\$/km)
Depreciação do veículo – 13.500,00	Manutenção – 2,25
Licenciamento e seguro – 2.025,00	Combustível – 1,00
Reserva para seguro – 1.600,00	Lubrificação – 0,05
Salários (motorista e ajudante) – 22.500,00	Pneus – 0,20

TABELA 1.2 Entrega mínima e máxima por dia (em toneladas)

	Segunda	Terça	Quarta	Quinta	Sexta
Entrega mínima	31	25	37	29	13
Entrega máxima	45	32	43	33	38

TABELA 1.3 Distância e demanda semanal de cada OM

OM	Demanda semanal (em t)	Distância (em km)
Alfa	8	12
Bravo	15	5
Charlie	16	140
Delta	18	160
Echo	20	15
Foxtrot	26	10
Golf	19	180
Hotel	14	06
India	16	25
Juliet	13	20

capacidade de transporte e gerar custos menores. Isto geralmente é possível quando o volume de carga é elevado, sendo mais econômico possuir um serviço de transporte próprio do que contratá-lo. Por outro lado, nem sempre isso é possível devido às exigências de requisitos especiais, tais como entrega rápida com confiabilidade muito elevada, indisponibilidade de algum equipamento especial, manuseio especial da carga ou um serviço que deva estar disponível quando necessário.¹

Na análise realizada neste capítulo, será considerado apenas o serviço de transporte por frota própria, em que os custos fixos e variáveis são perfeitamente conhecidos.

O gerente de tráfego de carga do DepSubMRJ é responsável pela decisão de utilização da frota, em que menores custos e melhor desempenho na entrega são perseguidos.

A montagem do plano de rota é um problema encontrado para direcionar os veículos através das diversas vias. O cálculo da rota pode ser feito pela mínima distância, pelo mínimo tempo ou por uma combinação de ambos. Além disso, o custo total é dependente não só do custo do transporte, mas também do custo de estocagem.²

Por outro lado, as empresas estocam produtos visando à melhoria da compensação entre oferta e demanda, além de diminuir seus custos totais. Neste sentido, os custos de armazenagem e manuseio de materiais podem ser justificados pela redução dos custos de transporte e de produção, pois os estoques podem reduzir os custos de transporte ao permitir que quantidades maiores e mais econômicas de carga sejam transportadas.

A questão que surge, então, é como obter, de forma determinística, a configuração ótima de carga e de trajeto da frota de caminhões, de modo a se obter o menor custo na distribuição de suprimentos da Marinha?

1. Ballou (2006).

2. Ballou (2008).

Para atacar o problema, foi usada a abordagem da Programação Linear Inteira, cujo arcabouço teórico está brevemente resumido a seguir.

PROGRAMAÇÃO LINEAR

A Programação Linear (PL) tem sido bastante utilizada desde a década de 1940 pelos militares e, até hoje, muitas empresas se beneficiam do seu uso. Segundo pesquisa da revista *Fortune*, 85% das empresas já utilizaram ou utilizam a Programação Linear. A PL teve como um de seus inventores George B. Dantzig, que descobriu o algoritmo Simplex, mecanismo matemático utilizado para resolver os problemas da PL, a qual serviu de base para outras ferramentas gerenciais tais como a Programação Não Linear, a Programação Inteira e outras técnicas de otimização. Outro matemático, Leonid G. Khanchian, descobriu o algoritmo de tempo polinomial, que teoricamente é superior a algoritmos polinomiais, como o algoritmo Simplex. Outros algoritmos mais rápidos foram descobertos mais tarde, como o algoritmo de Karmarkar.³

Todo modelo de Programação Linear tem duas características importantes: uma função-objetivo, a ser maximizada ou minimizada, e as restrições. Esse modelo de otimização restrito fornece meios de solucionar problemas em que os recursos são escassos de modo a otimizar um objetivo de interesse.⁴

O algoritmo Simplex utiliza uma série de definições e conceitos associados, entre os quais a definição de solução básica, o conceito de variável básica e não básica e a definição de solução básica viável. Uma solução viável é aquela que pode ser implementada considerando-se que os valores das variáveis de decisão estão de acordo com as restrições.⁵

O Simplex é o mesmo algoritmo utilizado nos softwares específicos para a solução de problemas de PL e caminha de uma solução viável para outra, diminuindo o valor da função-objetivo até o ponto ótimo ser alcançado. Esse algoritmo é simples e resolve uma gama de problemas e modelos do mundo real, gerando uma grande economia de recursos.⁶

Para m equações e n variáveis existem no máximo $n!/m!(n-m)!$ soluções básicas. Ainda assim, o método Simplex geralmente converge para a solução ótima em aproximadamente $3m/2$ iterações.⁷

Em relação às restrições da utilização do algoritmo Simplex, são conhecidos os problemas de escala quando existem muitos cálculos numéricos e a variabilidade entre o maior e o menor número é muito grande, causando problemas de cálculo numérico. Outros problemas dizem respeito às: (i) soluções inexistentes – quando não há solução que satisfaça as restrições; (ii) soluções ilimitadas – quando o algoritmo não consegue encontrar um ponto de máximo em um problema de maximização ou de mínimo num problema de minimização; e (iii) soluções múltiplas – quando mais de uma solução satisfaz todas as restrições e alcança um valor extremo da função-objetivo.⁸

3. Moore; Weatherford (2005).

4. Moore; Weatherford (2005).

5. Colin (2007).

6. Colin (2007).

7. Colin (2007).

8. Colin (2007).

O algoritmo Simplex não resolve qualquer problema de PL. Quando todas as variáveis são estritamente positivas nas soluções básicas viáveis, estamos diante de um problema conhecido como não degenerado. Em outras palavras, se uma ou mais das variáveis básicas numa solução têm o valor zero, a solução é denominada solução básica degenerada. A degeneração traz consigo problemas de convergência do algoritmo, sendo difícil identificar a solução ótima. Geralmente, os softwares utilizados no mercado, tais como o Solver, já tratam dessas particularidades, de forma invisível para o usuário.⁹

A Programação Inteira (PI) surgiu como um caso particular da PL, quando na resolução dos problemas havia a necessidade de uso de variáveis inteiras. O nível de dificuldade para resolução de problemas de PI é similar ao da resolução de problemas com PL, porém, enquanto estas possuem um espaço infinito de soluções, aquelas possuem um espaço finito de soluções, o que implica que a solução de PL seja mais fácil que a solução de PI.¹⁰

A PROGRAMAÇÃO DE VEÍCULOS SEM A PROGRAMAÇÃO INTEIRA

Atualmente, a distribuição da frota é realizada de modo empírico, baseada na experiência e competência dos profissionais envolvidos nessa tarefa. O objetivo da análise apresentada neste capítulo consiste em programar, de forma eficaz, a alocação de veículos frigoríficos do DepSubMRJ, atendendo às demandas semanais e minimizando os custos de distribuição. Existe uma programação semanal que atende tanto às OMs na Cidade do Rio de Janeiro quanto às cidades próximas. As entregas são realizadas de segunda-feira a sexta-feira, inclusive feriados. Como restrições, há uma programação de demanda e percursos em distância a percorrer para atender cada OM. A frota de veículos frigoríficos é constituída por caminhões de capacidades diferenciadas (14t, 8t e 3t), o que caracteriza a capacidade e a agilidade do caminhão no percurso. Os custos fixos e variáveis do transporte, a demanda semanal de cada OM e o valor máximo e mínimo de entrega diária levantados junto à instituição foram apresentados nas Tabelas 1.1 a 1.3).

A PROGRAMAÇÃO DE VEÍCULOS COM A PROGRAMAÇÃO INTEIRA

Para resolver a questão proposta foi utilizada a ferramenta Solver do Excel, buscando-se obter a programação que importasse menor custo possível. Como primeiro passo foi necessário identificar os fatores/variáveis que influenciavam diretamente no custo envolvido na distribuição de suprimentos da Marinha. A partir dessa primeira análise, chegou-se, então, a algumas premissas.

A primeira diz respeito à composição do custo total, que engloba o custo fixo, que não se altera, e o custo variável, que sofre influência de acordo com a programação executada. Neste caso, para diminuir o custo total, o que importa é fazer a distribuição

9. Colin (2007).

10. Colin (2007).

nas OM's com a menor quilometragem possível, pois é isso que influencia diretamente o custo variável.

O segundo pressuposto indica que, para se obter menor quilometragem, deve haver o mínimo possível de viaturas deslocando-se para as OM's, a fim de satisfazer suas necessidades. No caso em que haja a opção de suprir determinada OM, empregando uma ou duas viaturas, deve-se escolher a primeira alternativa, pois a segunda implica o dobro da quilometragem (um percurso por cada viatura).

A terceira premissa preconiza que é desejável suprir as OM's de uma só vez, pois, quanto mais viagens forem realizadas, maior será a quilometragem. As melhores escolhas são, portanto, aquelas em que se utiliza apenas um dia para suprir determinada OM.

Finalmente, deve-se observar que, embora não haja diferença no cálculo dos custos variáveis, deve haver penalização para a alocação de viaturas com pouca carga. Afinal, quando isto ocorre, está se deixando de aproveitar espaço útil nas viaturas, o que não deixa de ser uma perda, ainda que implícita.

Em seguida, foi realizada a programação no Excel, procurando-se encontrar uma resposta determinística para a questão de estudo, como alternativa para a programação da frota, executada apenas com base na experiência dos profissionais da Marinha.

Primeira tentativa de solução

Baseando-se nas premissas listadas anteriormente, buscou-se a princípio modelar o problema no Excel, criando uma tabela geral de variáveis ajustáveis, utilizando números inteiros, correspondendo à tonelagem alocada em cada viatura. Estas variáveis indicariam, diariamente, quais viaturas seriam empregadas, qual a capacidade utilizada das viaturas e que OM's seriam abastecidas.

Para um custo mínimo, ou seja, menor quilometragem, foram estabelecidas algumas restrições. Em primeiro lugar, diariamente, cada viatura não poderia ser abastecida com uma tonelagem acima de sua capacidade. Era imperativo também que as demandas de cada OM deveriam ser atendidas, ao término da semana. Por fim, as variáveis de decisão deveriam ser binárias. A Tabela 1.4 traz a configuração do problema para a segunda-feira, considerando-se que para cada dia da semana o número de células ajustáveis era a mesma. Dessa forma, o valor 0 significaria a não utilização da viatura, ao passo que o valor 1 indicaria o uso do veículo.

Apesar de a configuração do problema obedecer a uma lógica correta, ao se utilizar a ferramenta Solver do Excel não foi possível se chegar a uma solução, sendo exposta a seguinte mensagem: "Excesso de células ajustáveis". Sem dúvida, haveria grande necessidade de processamento dos cálculos para se testar todas as variáveis na tabela geral. Por este motivo, esta que seria a solução mais direta no Solver teve que ser descartada.

Em diversas outras tentativas de resolução do problema, as limitações de cálculo do Solver ficaram evidentes. Em alguns casos a execução do programa demorava horas sem que apresentasse uma solução. Como resultado dessas experiências, chegou-se à conclusão de que a modelagem do problema deveria considerar fases de resolução que não excedessem a capacidade da ferramenta empregada.

TABELA 1.4 Extrato das células variáveis da tabela geral

	Segunda						
	<i>OM A</i>	<i>OM B</i>	<i>OM C</i>	<i>OM D</i>	<i>OM E</i>	<i>OM F</i>	<i>OM G</i>
Vtr 1 (14t)	0	0	0	0	0	0	0
Vtr 2 (14t)	0	0	0	0	0	0	0
Vtr 3 (14t)	0	0	0	0	0	0	0
Vtr 4 (14t)	0	0	0	0	0	0	0
Vtr 5 (8t)	0	0	0	0	0	0	0
Vtr 6 (8t)	0	0	0	0	0	0	0
Vtr 7 (8t)	0	0	0	0	0	0	0
Vtr 8 (3t)	0	0	0	0	0	0	0
Vtr 9 (3t)	0	0	0	0	0	0	0
Totais	0	0	0	0	0	0	0

Primeira solução

Levando-se em consideração as premissas levantadas e a conclusão da necessidade de atacar o problema por fases, conforme anteriormente mencionado, chegou-se à primeira solução determinística com o uso da ferramenta Solver no Excel. Para essa primeira modelagem válida, alguns procedimentos foram adotados. Em primeiro lugar, foi criada uma planilha com parâmetros, para servir como base para as demais planilhas que seriam criadas. Além disso, a solução foi dividida em duas etapas, sendo empregada a ferramenta Solver em cada uma delas.

Na primeira etapa, para cada OM foi realizado um modelo Solver individual, com o propósito de verificar qual seria a melhor alocação de viaturas para satisfazer a demanda de uma única vez. A célula objetivo priorizou a menor quilometragem possível, com a menor ociosidade possível. Por este motivo, sua composição foi de viaturas empregadas, multiplicada pela distância, mais o espaço ocioso. Com isto, procurou-se tentar atender as restrições preestabelecidas.

Na segunda etapa, foi executado um Solver geral para decidir quais OMs seriam atendidas em cada dia da semana. Para tal, tomou-se como base o resultado da etapa anterior, que alocou a melhor quantidade de viaturas para satisfazer as OMs de uma única vez. Neste Solver foram impostas as restrições de quantidade de viatura e de que cada OM não poderia ser atendida mais de uma vez na semana. O mais importante na segunda etapa foi satisfazer as restrições. Por isso, a média foi um artifício utilizado para criar a variável objetivo; afinal, qualquer solução apresentaria o mesmo custo, pois a quilometragem seria sempre a mesma. Inicialmente, a variável-padrão foi definida,

tomando-se como base o desvio-padrão, porém, isso dificultou o trabalho do Excel para apresentar uma solução válida.

Cabe ressaltar que não foram utilizados números necessariamente binários para as variáveis de decisão na segunda etapa. No entanto, na verdade os números sempre eram 0 ou 1, devido à restrição de se atender a OM apenas uma vez na semana. As Tabelas 1.5 a 1.7 trazem o detalhamento de uma das configurações válidas encontradas nesta primeira solução determinística.

Segunda solução

A primeira solução não atendia o requisito de entregas máxima e mínima impostas pelo problema. Embora isso não fosse um prejuízo considerável, por não implicar maiores custos ou indisponibilidade de viatura, foi necessário elaborar uma segunda solução, que satisfizesse a todos os requisitos.

TABELA 1.5 OMs atendidas por dia

OM	Seg	Ter	Qua	Qui	Sex
Alfa	0	1	0	0	0
Bravo	0	0	1	0	0
Charlie	0	1	0	0	0
Delta	1	0	0	0	0
Echo	1	0	0	0	0
Foxtrot	0	1	0	0	0
Golf	1	0	0	0	0
Hotel	1	0	0	0	0
India	0	0	0	0	1
Juliet	0	0	0	1	0
Entrega	71	50	15	13	16

TABELA 1.6 Uso de viaturas

Uso Vtr	Seg	Ter	Qua	Qui	Sex
Vtr 14t	4	2	0	1	0
Vtr 8t	3	3	2	0	2
Vtr 3t	0	0	0	0	0

TABELA 1.7 Distribuição de viaturas e quilometragens envolvidas

OM	Vtr 14t	Vtr 8t	Vtr 3t	Demanda	Distância	Km por OM
Alfa	0	1	0	8	12	12
Bravo	0	2	0	15	5	10
Charlie	0	2	0	16	140	280
Delta	1	1	0	18	160	320
Echo	1	1	0	20	15	30
Foxtrot	2	0	0	26	10	20
Golf	1	1	0	19	180	360
Hotel	1	0	0	14	6	6
India	0	2	0	16	25	50
Juliet	1	0	0	13	20	20
Total						1.108

Cabe ressaltar que a segunda solução é mais trabalhosa que a primeira, embora seja a mais completa e a única que realmente satisfaz o problema proposto. Foram, portanto, adotados alguns procedimentos. Em primeiro lugar, criou-se uma planilha de parâmetros, de forma semelhante à anterior. Em seguida, dividiu-se a resolução em três fases.

Para a primeira fase partiu-se da premissa de que a melhor coisa a ser feita é suprir o máximo de OM's o mais rápido possível, obedecendo, é claro, as restrições de entrega máxima e mínima. Além disso, a OM deve ser suprida em apenas um único dia, para minimizar a quilometragem. Por este motivo, foi criado um Solver diário, para decidir quais OM's seriam supridas diariamente.

A variável-objetivo foi a sobra de toneladas a serem supridas, devendo ser a mínima possível. Cabe ressaltar que as planilhas estavam relacionadas de forma que a demanda considerada em determinado dia fosse igual à demanda que deixou de ser atendida no dia anterior, de modo a não se atender mais de uma vez a determinada localidade. A Tabela 1.8 traz o resultado do Solver para segunda-feira.

Com os dados da primeira fase chegou-se a uma tabela que demonstrava como as OM's seriam atendidas, ao longo da semana, conforme exposto na Tabela 1.9. Faltava, no entanto, determinar como fazer a distribuição das viaturas para que fosse atendido o objetivo maior do problema: determinar a melhor programação para reduzir custos. Para tal, foi realizada a segunda fase.

Como base para a segunda fase, além da Tabela 1.9, foram adotados na variável-objetivo os princípios de menor quilometragem possível e também o de menor capacidade ociosa possível (espaço vazio em viatura utilizada). Estes são os mesmos parâmetros utilizados na solução anterior.

TABELA 1.8 Primeira fase (segunda-feira)

Entrega (máx)	45	45		
Entrega (mín)	31	45	Efetuada	Resta
Alfa	8	0	0	8
Bravo	15	0	0	15
Charlie	16	0	0	16
Delta	18	0	0	18
Echo	20	0	0	20
Foxtrot	26	1	26	0
Golf	19	1	19	0
Hotel	14	0	0	14
India	16	0	0	16
Juliet	13	0	0	13
Total	165			120

TABELA 1.9 Consolidação da primeira fase

OM	Seg	Ter	Qua	Qui	Sex	Dist
Alfa	0	0	1	0	0	12
Bravo	0	0	1	0	0	5
Charlie	0	0	0	1	0	140
Delta	0	1	0	0	0	160
Echo	0	0	1	0	0	15
Foxtrot	1	0	0	0	0	10
Golf	1	0	0	0	0	180
Hotel	0	1	0	0	0	6
India	0	0	0	1	0	25
Juliet	0	0	0	0	1	20
Totais	45	32	43	32	13	

TABELA 1.10 Segunda fase (segunda-feira)

	Vtr (t)		Dist envo				
	14	8	3		Demanda	Dem at	Dem ref
	0	0	0	0	8	0	0
	0	0	0	0	15	0	0
	0	0	0	0	16	0	0
	0	0	0	0	18	0	0
	0	0	0	0	20	0	0
	2	0	0	20	26	28	26
	1	1	0	360	19	22	19
	0	0	0	0	14	0	0
	0	0	0	0	16	0	0
	0	0	0	0	13	0	0
Utl	3	1	0	<=	Cap total	50	45
Disp	4	3	2				

As restrições adotadas em cada Solver desta segunda fase referiam-se, sobretudo, à limitação de viatura. A **Tabela 1.10** traz o resultado do Solver de segunda-feira, nesta segunda fase: “Dem at” refere-se à capacidade das viaturas empregadas, enquanto “Dem ref” diz respeito à necessidade das OM.

Na última fase dessa resolução, foi feita apenas a elaboração de uma planilha de consolidação, com o resultado final da primeira e da segunda fases da resolução. A terceira fase detalha, portanto, diariamente, o percurso das viaturas que serão empregadas na distribuição de suprimentos pelas OM.

Comparação entre as soluções

A única vantagem da primeira solução em relação à segunda é ser relativamente mais simples, exigindo menos tabelas Solver. Esta aparente vantagem, no entanto, talvez nem mesmo possa ser considerada, tendo em vista que somente na segunda solução foram atendidos os requisitos propostos pelo problema de entrega máxima e mínima.

Além disso, a segunda solução utiliza de modo mais eficaz o espaço das viaturas. Isto acontece porque ela não considera as OM de forma isolada, e sim a entrega que deve ser efetuada no dia. Evita-se, com isso, a sobreposição de sobras de espaços vazios nas viaturas quando estas são alocadas, considerando-se as OM de forma isolada. Como resultado, pode-se, inclusive, utilizar viaturas de três toneladas, o que não ocorria na primeira solução.

Finalmente, é possível concluir que a segunda solução é, com segurança, a ótima, porque está em consonância com o objetivo maior da ferramenta Solver, que é apontar a melhor solução. No caso da primeira solução, podem existir alternativas melhores, e se assim for o instrumento de apoio à decisão está comprometido. Por isso, embora mais trabalhosa, a segunda solução é a que deve ser considerada na solução do problema proposto.

CONSIDERAÇÕES FINAIS

A resolução do problema real com o uso do Excel permitiu que se chegasse a uma visão acerca das potencialidades e das limitações da ferramenta Solver. Verificou-se, nesse sentido, a necessidade de executar a modelagem do problema, considerando o processamento eletrônico que possui limitações. No entanto, após tratar o problema em fases, a ferramenta Solver pode resolver questões complexas, como a considerada neste trabalho.

Em relação às limitações da ferramenta Solver reportadas na literatura, tomou-se cuidado em relação ao problema mencionado de escala para que não houvesse muita variabilidade entre o maior e o menor número nos cálculos efetuados, o que poderia gerar problemas de convergência.¹¹

Outra limitação da ferramenta Solver que deve ser mencionada diz respeito ao problema da degeneração. Conforme reportado na literatura,¹² se uma ou mais das variáveis básicas numa solução tem o valor zero, a solução é chamada solução básica degenerada, e isto implica problemas de convergência do algoritmo, sendo difícil identificar a solução ótima. Ainda que o algoritmo Simplex trate esses problemas, este fator deve ser considerado na tentativa de resolução na primeira solução.

Entretanto, na primeira solução, para atender os requisitos preestabelecidos e apresentar um valor para a célula alvo que fosse ótimo, o Solver apresentou, inicialmente, em alguns casos, uma demora excessiva. Conforme reportado na literatura,¹³ para m equações e n variáveis deveria haver, no máximo, $n!/m!(n-m)!$ soluções básicas e o algoritmo deveria convergir para a solução ótima em aproximadamente $3m/2$ iterações. Trazendo essas fórmulas para o problema em questão, verifica-se que, no primeiro caso, existiam 362 equações relacionadas às restrições e 315 variáveis, o que poderia levar a 543 iterações, explicando a demora na solução e a não convergência para um valor ótimo. Por outro lado, com a solução vislumbrada, utilizando-se a defasagem dos dias da semana, conseguiu-se uma convergência com 18 iterações.

Durante a resolução do problema, constatou-se que a formulação de um bom modelo é crucial para a sua solução e que é importante a análise crítica detalhada para avaliação da real necessidade de se modelar uma variável como inteira. Tal constatação foi baseada na falta de convergência do resultado na primeira solução e no sucesso alcançado com a nova modelagem na segunda solução, considerando-se a defasagem proposta. Cabe considerar que, embora trabalhosa a montagem do problema, após os

11. Colin (2007).

12. Colin (2007).

13. Colin (2007).

dados serem correlacionados na planilha não existem dificuldades para se efetuar novas soluções a partir de modificações que venham a ocorrer. No caso, por exemplo, de se acrescentarem viaturas ou haver aumento de demanda, a partir da mudança na planilha de parâmetros todos os demais cálculos serão efetuados de forma automática. A solução apresentada não serve, portanto, apenas para responder ao momento atual, podendo ser utilizada também em ocasiões futuras.

Outro aspecto importante que deve ser salientado é que nas soluções de problemas reais desta natureza não devem ser considerados apenas os custos e parâmetros explícitos. Trafegar com viaturas com espaço ocioso, matematicamente, não apresentaria aumento de custos. Conforme mencionado na literatura,¹⁴ na logística o custo total é dependente não só da variável custo do transporte, mas também do custo de estocagem. No entanto, sabe-se que, ao menos indiretamente, isto representa perda, pelo espaço ocioso. De forma semelhante, diante de um mesmo custo, é melhor atender o mais rápido possível a demanda do usuário. Cabe, portanto, a quem vai modelar o problema, a tarefa de considerar estes e outros aspectos implícitos na resolução de questões reais.

Como indicação para pesquisas futuras, fica a sugestão de se tentar modelar tais problemas de modo ainda mais realista, considerando todos os aspectos mais importantes em relação à situação em foco. Por exemplo, a demanda de cada OM poderia ser considerada de forma probabilística, através de uma distribuição de probabilidades condizente com o seu histórico de comportamento, o que estaria mais de acordo com a realidade. Isso seria possível por meio do uso de uma abordagem probabilística, como a Simulação de Monte Carlo,¹⁵ em vez da Programação Inteira, utilizada neste trabalho, que considerou a demanda de forma puramente determinística.

Cabe ressaltar, no entanto, que, no caso de modelagens ainda mais realistas (e, portanto, potencialmente mais complexas), tornam-se ainda mais importantes os cuidados em relação à modelagem mencionados nestas considerações finais, para que o esforço computacional na busca da solução ótima não ultrapasse os limites da viabilidade operacional.

EXERCÍCIOS

1. Uma empresa tem que planejar seus gastos em Pesquisa e Desenvolvimento para os próximos cinco anos. A empresa pré-selecionou quatro projetos e deve escolher dentre estes quais priorizar. Os dados relevantes ao problema encontram-se na planilha **Exercício 1**. Nela também se encontra a disponibilidade de capital a ser alocado em cada um dos anos, bem como o valor presente líquido (VPL) de cada projeto. Como todos os projetos apresentam VPL positivos, todos seriam candidatos. Vale notar que existe uma limitação no valor a ser investido anualmente.
 - a. Quais projetos devem ser priorizados?
 - b. Agora considere que os projetos 1, 3 e 4 são concorrentes e apenas um deles pode ser escolhido. Quais projetos devem ser priorizados?

14. Ballou (2008).

15. Hertz (1980); Paula *et al.* (2007).

- c. Sem considerar a letra b, dessa vez considere que o projeto 1 depende de uma tecnologia que deve ser desenvolvida pelo projeto 3, isto é, só pode ser aprovado se o projeto 3 também o for. Quais projetos devem ser priorizados?
2. Um cliente procurou um analista de investimentos com o objetivo de investir em ações. O analista o entrevista para identificar o seu perfil de risco e apresentar as opções de investimento. O cliente está decidido a investir em papéis de bancos. O analista deve apresentar uma proposta de investimento que atenda às expectativas do cliente. O retorno e o risco de cada investimento estão apresentados na planilha. **Exercício 2.** Durante a entrevista, o cliente fez as seguintes exigências:
- pelo menos um quarto da carteira deve ser composta por ações do Banco A, devido ao seu baixo risco;
 - a participação das ações do Banco B deve ser de, no máximo, 12,5%, devido ao seu alto risco;
 - a participação de Renda Fixa na carteira deve ser de, no máximo, 10%, devido ao seu baixo retorno;
 - a taxa máxima suportada de risco é de 5,5%.
3. Uma corretora de valores gerencia recursos de terceiros através da composição de carteiras de investimento para clientes, com diferentes perfis de risco. Um deles exige que:
- não mais de 30% do total aplicado seja aplicado em um único investimento;
 - um valor não inferior a 40% do total aplicado seja investido em títulos de maturidade maior do que 10 anos;
 - o total aplicado em títulos de alto risco seja, no máximo, de 40% do total investido.
- A planilha **Exercício 3** mostra os dados dos títulos selecionados. Determine qual percentual do total deve ser aplicado em cada título.
4. Um indivíduo dispõe de R\$ 6.000 para investir em até dois ativos. Ele tem três opções de ativos, que vão fornecer o retorno após cinco anos. Cada um dos investimentos pode ser feito em valores múltiplos de R\$ 1.000 e ele deseja investir, no máximo, R\$ 4.000 em cada ativo. Os retornos esperados apresentam ganhos de escala (ou seja, sempre é melhor fazer uma aplicação de R\$ 3.000 em um ativo do que uma aplicação de R\$ 1.000 e outra de R\$ 2.000 no mesmo ativo, por exemplo) e podem ser visualizados na planilha **Exercício 4**. Qual deve ser a política de investimento?
5. Uma empresa possui as seguintes opções de investimento em projetos para os próximos quatro anos: investimento maciço em P&D, política agressiva de Vendas, abertura de uma nova loja e recapacitação dos Recursos Humanos. Enfrentando limitações anuais de capital, a empresa deve escolher em quais projetos investir para obter o maior retorno possível, a valor presente. O VPL dos projetos, o capital necessário para cada um deles e a projeção do capital disponível estão ilustrados na planilha **Exercício 5**.

- a. Determine que projetos a companhia deve selecionar para maximizar o valor presente líquido do capital investido.
- b. Considere que, devido a questões estratégicas, a empresa decida que a recapacitação dos Recursos Humanos é um projeto que não pode deixar de ser realizado, independentemente de seu VPL estimado. Que mudanças esta informação gera na formulação do problema e no resultado?
- c. Sem considerar o que foi mencionado na letra b, o que você recomendaria se a companhia conseguir um adicional de R\$ 10.000 para cada um dos quatro anos?
- d. E considerando o que foi mencionado na letra b?

REFERÊNCIAS

- BALLOU, R. *Logística empresarial*. São Paulo: Atlas. 2008.
- _____. *Gerenciamento da cadeia de suprimentos/Logística empresarial*. Porto Alegre: Bookman, 2006.
- COLIN, E. *Pesquisa operacional*. Rio de Janeiro: LTC. 2007.
- HERTZ, D. *Análise de risco em investimentos de capital*. *Biblioteca Harvard de Administração de Empresas*, 8 (3), p. 1-14, 1980.
- MOORE, J.; WEATHERFORD, L. *Tomada de decisão em administração com planilhas eletrônicas*. Porto Alegre: Bookman. 2005.
- PAULA, R.; CAPELO JR. E.; COSTA, C. O Cálculo do Valor Presente Líquido com Tratamento do Risco através do Método de Simulação de Monte Carlo. In: Encontro da ANPAD, 31, 2007, Rio de Janeiro. Anais... Rio de Janeiro: ENANPAD. 2007.

Avaliação da escolha de um fornecedor sob condição de riscos a partir do método de Árvore de Decisão

Wesley Vieira da Silva
PPAD/PUC-PR

Jansen Maia Del Corso
PPAD/PUC-PR

Hanna Tatarchenko Welgacz
PPAD/PUC-PR

Julinês Bega Peixe
PPAD/PUC-PR

Objetivos de aprendizagem do capítulo

- Saber modelar um problema real em situação de risco através de uma Árvore de Decisão.
- Entender o conceito e as diferenças entre nó de decisão e de evento.
- Compreender o conceito do Valor Monetário Esperado.
- Aprender a estimar o Valor da Informação Imperfeita e o Valor da Informação Perfeita.
- Saber conduzir uma Análise de Sensibilidade acerca dos parâmetros mais importantes.
- Conseguir considerar, na análise, o perfil de aversão a risco do tomador de decisão.

INTRODUÇÃO

Uma decisão precisa ser tomada quando se está diante de um problema que dispõe de mais de uma alternativa para sua solução.¹ A Teoria da Decisão parte do pressuposto de que os indivíduos são capazes de expressar suas preferências básicas, e são racionais, quando enfrentam situações de decisões simples. Com base nessa proposição, a metodologia desenvolvida pela Teoria da Decisão possibilita a resolução de problemas

1. Gomes; Gomes; Almeida (2002).

de decisão mais complexos. Estes problemas são comuns em uma infinidade de áreas, tanto públicas quanto privadas, e desde tempos remotos o homem tenta resolvê-los apoiando-se em abstrações, heurística e raciocínios dedutivos, a fim de validar e guiar escolhas. Para resolver tais problemas, até a primeira metade do século XX, utilizava-se basicamente a esperança matemática para tomada de decisão em condições aleatórias; porém, em muitas situações, o risco associado a tal procedimento era inaceitável.

No processo de modelagem da realidade, o pesquisador troca a riqueza e a abrangência da realidade pelo poder de análise e pela capacidade de experimentação. Neste caso, os elementos do modelo a ser formulado podem ser vistos como variáveis de controle ou de decisão e ser manipulados para se atingir os objetivos. O modelo compreende ainda variáveis de estado ou de natureza, sobre as quais não se tem controle, mas que afetam as consequências ou resultados da decisão.

O ambiente atual dentro e fora das organizações é dinâmico e cheio de incerteza. Os gerentes estão sobrecarregados devido à complexidade dos processos organizacionais. As exigências diárias em relação a decisões rápidas e assertivas demandam o uso de ferramentas de tomada de decisão que possam aumentar a qualidade e a velocidade nesse processo.

O processo de tomada de decisão dentro das organizações modernas demanda um profissional que conheça o ambiente no qual a empresa está inserida e que também esteja ciente das suas constantes e frequentes mudanças.

Em constante cobrança por eficiência, o profissional deve dispor de conhecimentos técnicos sobre o assunto em relação ao qual vai opinar e também ser ágil no levantamento de todas as alternativas possíveis. Para atender essa necessidade, o uso dos métodos quantitativos evoluiu para o conceito da Teoria da Decisão, dentro da qual se encontram diversas ferramentas e técnicas que contribuem para maior eficácia gerencial. Os métodos quantitativos modernos desenvolveram a abordagem lógica para situações de incerteza e risco, o que se aplica ao ambiente atual das organizações.

É diante de uma dessas situações de risco que se pode recorrer a uma ferramenta que permita, a partir de probabilidades conhecidas ou subjetivamente atribuídas e com o auxílio de heurísticas, visualizar e interpretar facilmente algumas opções para a tomada de decisão mais acertada.

ÁRVORE DE DECISÃO

Na década de 1970, começaram a surgir os primeiros métodos de apoio com o intuito de enfrentar situações específicas, nas quais um tomador de decisão, atuando com racionalidade, deveria resolver um problema em que vários eram os objetivos a serem alcançados de forma simultânea. Daí surgiu a Árvore de Decisão como um dos mais importantes métodos de análise de decisão, pautando-se em conceitos como o de risco e retorno.

As Árvores de Decisão utilizam a estratégia *dividir-e-conquistar* (*divide-and-conquer*), em que árvores são construídas utilizando-se apenas alguns atributos. As Árvores de Decisão são uma das técnicas de aprendizado de máquina (*machine learning*), na qual um problema complexo é decomposto em subproblemas mais simples. Recursivamente, a mesma estratégia é aplicada a cada subproblema.²

2. Gama (2004).

Quinlan, professor da Universidade de Sidney, considerado o “pai das Árvores de Decisão”, elaborou um novo algoritmo, chamado *ID3*, desenvolvido em 1983. O *ID3* e suas evoluções (*ID4*, *ID6*, *C4.5*, *See 5*) são algoritmos muito utilizados para gerar Árvores de Decisão. O atributo mais importante é apresentado na árvore como o primeiro nó, e os atributos menos importantes, segundo o critério utilizado, são mostrados nos nós subsequentes. As principais vantagens das Árvores de Decisão são o fato de que elas “tomam decisões” levando em consideração os atributos que são considerados mais relevantes, segundo a métrica escolhida, além de serem compreensíveis para as pessoas. Ao escolher e apresentar os atributos em ordem de importância, as Árvores de Decisão permitem aos usuários conhecer quais fatores mais influenciam os seus trabalhos.

Dentro do conceito do método de Árvore de Decisão está o modelo de árvore binomial, com o qual se pode descrever uma multiplicidade de alternativas numa escolha de opção otimizada. A tarefa deste método é estabelecer uma relação entre a situação de tomada de decisão e uma expressão numérica das opções disponíveis.

Um problema representa o início de uma árvore que se compõe de ramos, nós e folhas. Para cada nó deve ser definido o próximo critério de ramificação, uma decisão ou cenário. Os ramos representam decisões e cenários, e as folhas são alternativas finais para a escolha de uma única e melhor opção. Ao receber novos dados, a árvore modifica-se e gera uma nova decisão otimizada.

O problema em construir uma Árvore de Decisão pode ser expresso recursivamente: primeiro deve-se selecionar um atributo para colocar no nó-raiz e fazer um ramo para cada possível valor. Isso divide o problema em subconjuntos, um para cada valor do atributo. Agora o processo pode ser repetido recursivamente para cada ramo. Se a qualquer instante todos os exemplos em um nó têm a mesma classificação, interrompe-se o desenvolvimento daquela parte da árvore. Como determinar, no entanto, qual atributo dividir? Escolhe-se o atributo que gere uma árvore menor e que tenha chances de classificar melhor, ou seja, precisa-se mensurar o grau de pureza de cada nó. Com isto, é possível escolher o atributo que produz os nós filhos mais puros.³

A Árvore de Decisão está diretamente ligada com o que se denomina flexibilidade gerencial, pois o método de análise permite rever a estratégia inicial e alterar planos de acordo com novas condições econômicas. As flexibilidades gerenciais possibilitam tanto capitalizar futuras oportunidades favoráveis ao negócio quanto diminuir perdas, isto é, melhoram o potencial de ganhos e limitam o potencial de perdas.⁴ A flexibilidade gerencial consiste em um conjunto de opções reais tais como postergar um projeto, expandir ou contrair escala de produção ou avaliar a alteração das matérias-primas de um projeto.

Considera-se a flexibilidade gerencial como uma possibilidade, não como uma obrigação. Muitos autores acreditam que as Opções Reais no futuro não constituirão meramente uma ferramenta analítica mas, em contextos mais rigorosos e intensivos, um processo organizacional. Contudo, a Árvore de Decisão é um método que permite indicar, de forma gráfica e cronológica, um caminho a ser seguido em um processo de decisão, explicitando etapas a serem cumpridas para alcançar o objetivo pretendido.

3. Carvalho (2002).

4. Minardi (2004).

A árvore representa um processo de decisão em que os nós são os momentos no tempo em que o decisor deve efetuar uma decisão. As etapas da técnica para o auxílio da tomada de decisão são definidas como: a) definição do tema; b) definição do objetivo, metas e submetas; c) construção da Árvore de Decisão; d) revisão da Árvore de Decisão; e) encerramento.⁵

Algumas vantagens da utilização desta técnica podem ser relacionadas:⁶

- permite a subdivisão do objetivo em metas e submetas, indicando como alcançá-las;
- orienta o tomador de decisão à medida que este responde à pergunta: *O que é necessário fazer para alcançar a meta pretendida?*;
- permite o exame, pelo tomador de decisão, de todas as possibilidades;
- permite a criação de algoritmos facilmente implementados em computadores.

Uma decisão tomada com base em todos os dados e informações disponíveis e alternativas possíveis e que, principalmente, utilize ferramentas de métodos quantitativos apropriados, deve ser considerada como uma boa decisão, mesmo que apresente algo que um administrador não vise como meta. Complementando essa convicção, uma decisão que não considerou todos os procedimentos prescritos pela metodologia é considerada má decisão, mesmo que o resultado obtido seja favorável. Isso porque, como não foram exploradas todas as informações e alternativas possíveis, não se pode afirmar que o resultado favorável foi otimizado.⁷

Vale destacar que as Árvores de Decisão garantem maior flexibilidade do que os mecanismos tradicionais na medida em que as decisões são definidas, tal como fora mencionado anteriormente, com maior grau de flexibilidade. Nas árvores, uma série de eventos podem ser mapeados ao longo dos diversos ramos, envolvendo várias decisões de sequenciamento. A Figura 2.1 apresenta uma Árvore de Decisão simplificada.

Constata-se com base na Figura 2.1 que os valores de “p” referem-se às probabilidades de ocorrência de um evento. A Árvore de Decisão incorpora conceitos de incerteza na análise de uma alternativa de investimento ou mesmo de financiamentos, o que a torna uma técnica superior às técnicas tradicionais.

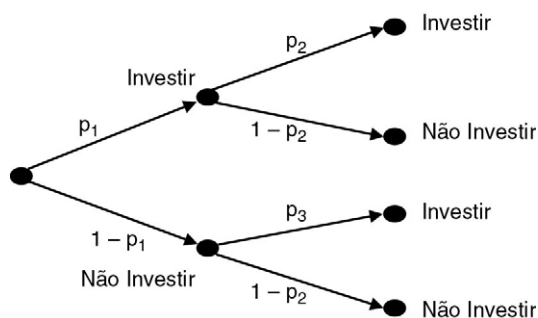


FIGURA 2.1 Exemplo de uma árvore de decisão.

5. Gomes; Gomes; Almeida (2009).

6. Gomes; Gomes; Almeida (2009).

7. Corrar; Theóphilo (2004).

A ESCOLHA DE UM FORNECEDOR COM O AUXÍLIO DA ÁRVORE DE DECISÃO

Diante das necessidades operacionais de um gerente da importadora brasileira Metalúrgica e Distribuidora de Aço, o exemplo a seguir ordena diretamente a utilização de uma ferramenta que visa auxiliá-lo na tomada de decisão. O objetivo é determinar a melhor escolha entre três fornecedores de aço para importação pautando-se na técnica da Árvore de Decisão. A determinação do fornecedor torna-se muito importante porque existem diversas variáveis que, caso não sejam cumpridas, vão comprometer a imagem da empresa distribuidora perante o mercado. As variáveis não são somente os custos, mas também o prazo de entrega, a qualidade de produto para aceitação do mercado importador e outros, que são fatores determinantes na escolha do fornecedor.

A Árvore de Decisão é um método dinâmico, fácil de utilizar e pode ser aproveitado para futuras consultas, criando assim um banco de dados das decisões. O desafio é aproximar os dados descritivos a uma análise quantitativa e evitar um julgamento intuitivo na tomada de decisão.

A análise do processo de decisão se desdobra em duas etapas: a primeira é a decisão de importar ou não, e a segunda refere-se à escolha entre três fornecedores estrangeiros, com exploração prévia de atributos de avaliação.

A Teoria da Decisão estabelece procedimentos para um processo de tomada de decisão. Os passos a seguir analisam o problema proposto à luz dos conceitos desse processo, procurando, de forma didática, determinar a melhor alternativa a ser seguida.⁸

Passo 1: Definição do problema

A Metalúrgica e Distribuidora de Aço é uma empresa brasileira de grande porte que usa como insumo tarugos, chapas e bobinas de aço para produção de tubos, cantoneiras, perfis e outros produtos; além disso, é distribuidora desses produtos para metalúrgicas de menor porte. A diretoria da empresa observou a tendência de aquecimento da demanda para seus produtos devido a alguns fatores.

Os produtos são comprados de fabricantes de aço nacionais que, devido à grande demanda para consumo próprio, suprem as necessidades das metalúrgicas a partir das quotas preestabelecidas. Dessa maneira fica impossível aumentar os volumes de compra de imediato. Caso o aquecimento da demanda se concretize, tal como indicam várias pesquisas das associações comerciais do setor metalúrgico brasileiro, a Metalúrgica e Distribuidora de Aço precisará aumentar seu estoque, devido à sua necessidade, passando de 1 mil toneladas por mês para 4 mil toneladas para o mesmo período. O problema mais evidente no primeiro momento é onde obter a quantidade de insumos além da quota com fornecedor nacional.

Passo 2: Alternativas de ação

Sendo impossível obter o volume adicional de 3 mil toneladas de aço no mercado nacional, a diretoria tem algumas alternativas, que se alinham em duas etapas de

8. Corrar; Theóphilo (2004).

tomada de decisão, sendo a segunda somente necessária caso opte positivamente pela primeira alternativa.

Primeira etapa de decisão:

1. importar o produto do exterior, avaliando com otimismo a situação do mercado em expansão; ou
2. permanecer no mesmo volume de 1 mil toneladas mensais, sem recorrer à compra adicional via importação, devido às incertezas de crescimento do mercado de construção.

Segunda etapa de decisão:

1. Caso seja tomada a decisão de importar, deve-se escolher um fornecedor estrangeiro dentro de várias ofertas.

O desafio do profissional que toma a decisão é precisar ou obter dados quanto à probabilidade de ocorrência desse cenário. Essas probabilidades são estimadas em termos percentuais. Entre os fatores que identificam o ambiente de incerteza na tomada de decisão do problema estudado estão aumento da demanda no mercado interno, mudanças na política industrial no Brasil ou no país exportador e volatilidade da taxa de câmbio de dólar norte-americano, entre outros.

Baseando-se em dados do mercado, o profissional conseguiu atribuir as probabilidades de ocorrência de cenários, concluindo que a decisão seja tomada em um ambiente de risco. A condição do cenário de aquecimento do mercado configura um exemplo de opção real, em que a decisão de importar pode ser tomada após a revelação da informação. Mas, no caso analisado, o distribuidor deseja estar preparado para o aumento da demanda devido aos prazos extensos de entrega da mercadoria na importação e para estar com o produto à disposição de compradores.

Para o tomador da decisão, a situação ideal seria aquela em que a probabilidade de ocorrer um cenário fosse estimada em 100%. Nessa situação, pode-se considerar que o tomador de decisão está diante de uma informação perfeita, que transformaria um ambiente de risco em um ambiente de certeza. Com base em informações sobre mercado previamente coletadas, o dirigente admite que as probabilidades de ocorrência dos cenários de primeira etapa sejam consideradas como as da [Tabela 2.1](#).

Em se tratando de recusa à importação de um estoque adicional, não há necessidade de próxima etapa de decisão, ou seja, de segundo nível. Mas, ao optar pela importação do produto, a empresa encontra alguns riscos conjunturais do país exportador e riscos

TABELA 2.1 Cenários e suas probabilidades de ocorrência

Cenários para a primeira etapa de decisão: Importar	Probabilidades
Cenário ₁ (C ₁): Aquecimento do mercado	80%
Cenário ₂ (C ₂): Mercado Estável	20%

próprios do negócio a ser efetuado. Também existe um grau de incerteza nas avaliações feitas em relação aos fornecedores, o que significa que um fornecedor apenas bom pode ter sido avaliado como ótimo e vice-versa.

Um cenário de risco conjuntural presente no trabalho é a situação econômica do setor siderúrgico na China. Devido à crescente produção e demanda do aço naquele país, o governo pode intervir a qualquer momento, limitando a produção de aço nacional com o intuito de não desvalorizar o preço das exportações chinesas. A probabilidade atribuída pelo profissional à ocorrência dessa ação é de 60%.

Portanto, antes de optar pelo fornecedor chinês, deve-se observar a probabilidade de que este, diante de novas políticas industriais, poderá não cumprir o contrato de venda, e o importador brasileiro ficará sem o produto para a data acordada. Tal ocorrência também deve ser analisada do ponto de vista financeiro, mesmo não incorrendo em gastos, pois pode acarretar custos como colocação de pedido, abertura de carta de crédito e, ainda, o prejuízo à imagem da empresa por não dispor do produto conforme demanda o mercado, o que representa uma limitação para esse estudo, já que não será levada em conta.

O risco conjuntural para fornecedores franceses e ucranianos deve ser estimado em 10%, pois sempre há uma margem de risco. Assim, baseando-se na probabilidade de ocorrência de intervenção do governo, a Tabela 2.2 mostra a probabilidade de sucesso e insucesso na entrega de mercadoria.

Passo 3: Possíveis resultados decorrentes das alternativas

Cada alternativa de ação terá um ganho e uma perda determinados pelo sucesso ou insucesso da decisão. A margem de contribuição estabelecida, *a priori*, é de 16,6%. Uma estimativa de venda do produto do fornecedor nacional acaba resultando nas seguintes operações matemáticas:

Quantidade	3.000 toneladas
Receita Líquida de Vendas (RLV):	US\$600,00/t * 3.000/t = US\$1.800.000,00
Custo das Mercadorias Vendidas (CMV):	US\$500,00/t. * 3.000/t. = US\$1.500.000,00
Margem de contribuição (MC):	US\$1.800.000,00 – US\$1.500.000,00 = US\$300.000,00
Margem de contribuição (%):	US\$300.000,00/US\$1.800.000,00 = 16,6%

TABELA 2.2 Probabilidades de sucesso e insucesso na entrega de mercadoria

Cenários para a segunda etapa de decisão: escolha do fornecedor estrangeiro em função do risco conjuntural	Probabilidade de sucesso	Probabilidade de insucesso
Cenário ₁ (c ₁): Fornecedor francês	90%	10%
Cenário ₂ (c ₂): Fornecedor chinês	40%	60%
Cenário ₃ (c ₃): Fornecedor ucraniano	90%	10%

TABELA 2.3 Possíveis resultados decorrentes das alternativas em função do mercado

Decisão	Lucro/Prejuízo C ₁ : Aquecimento do mercado	Lucro/Prejuízo C ₂ : Mercado estável
Importar	Margem de contribuição 16,6% (US\$300.000,00)	Retorno do investimento em quatro meses; Perda de aplicação no mercado financeiro do capital investido em estoque importado por três meses (até uso completo do estoque importado): 1,2% a.m. em três meses = $(1 + 0,012)^3 = 3,6\%$ de perda de oportunidade de investimento.
Não importar	Taxa de retorno financeiro de aplicação do capital por um mês, 1,2% a.m.; Perda de oportunidade de investimento de ganho de 16,6%.	Taxa de retorno financeiro de aplicação do capital por um mês, 1,2% a.m.
Probabilidade de ocorrência do cenário	80%	20%

Por outro lado, caso a empresa opte por não importar, ela poderá aplicar seu capital em títulos públicos e ter uma remuneração média definida, *a priori*, de 1,2% a.m. O resumo das alternativas está evidenciado na [Tabela 2.3](#).

Passo 4: Listar o retorno de cada combinação de alternativas e consequências (Tabela *Payoff*)

O resultado de cada decisão é um valor, *payoff*, como é conhecido na Teoria da Decisão, que é o valor presente dos lucros dos diversos períodos futuros envolvidos. A tabela que evidencia os *payoffs* foi desenvolvida pautando-se nas seguintes operações matemáticas:

Margem de contribuição na decisão de importar:

1. Importar em cenário de aquecimento do mercado

$$\text{US\$1.800.000,00} - \text{US\$1.500.000,00} = \text{US\$300.000,00}.$$

2. Importar em cenário de mercado estável

Margem de contribuição: US\$300.000,00.

Perda de aplicação no mercado financeiro: $\text{US\$1.500.000,00} * (1 + 0,012)^3 \approx \text{US\$54.000,00}$

Margem de contribuição esperada: $\text{US\$300.000,00} - \text{US\$54.000,00} \approx \text{US\$246.000,00}$

O valor monetário esperado da decisão de importar é obtido por meio da multiplicação das probabilidades de ocorrência de cada cenário por valores de margem de contribuição, tal como a seguir:

$$\begin{aligned} & \text{US\$}300.000,00 \times 0,80 + \text{US\$}246.000,00 \times 0,20 \\ & = \text{US\$}289.000,00. \end{aligned}$$

Margem de contribuição na decisão de não importar:

1. Não importar em cenário de aquecimento do mercado

$$\text{US\$}1.500.000,00 \times 1,2\% = \text{US\$}18.000,00.$$

2. Não importar em cenário de mercado estável

$$\text{US\$}1.500.000,00 \times 1,2\% = \text{US\$}18.000,00.$$

O valor monetário esperado da decisão de não importar é obtido a partir da multiplicação das probabilidades de ocorrência de cada cenário por valores de margem de contribuição, tal como se demonstra a seguir:

$$\text{US\$}18.000,00 \times 0,80 + \text{US\$}18.000,00 \times 0,20 = \text{US\$}18.000,00.$$

Na Tabela 2.4, vê-se um resumo das operações matemáticas realizadas anteriormente, considerando as probabilidades associadas às decisões a serem tomadas.

Segundo o método de análise dos resultados da tabela de *payoff*, designado anteriormente, deve-se escolher a decisão que apresente maior Valor Monetário Esperado (VME), ou seja, deve-se decidir pela importação, pois o maior VME apresentado é de US\$289.200,00.

Passo 5: Selecionar um dos modelos utilizados na Teoria da Decisão, aplicá-lo e tomar a decisão

Ao analisar as variáveis e o ambiente do problema, foi escolhido o método de Árvore de Decisão devido ao fato de trabalhar com fatores de análise em um ambiente de risco, com atribuição de probabilidade de ocorrência de certas condições do mercado.

TABELA 2.4 Valor monetário esperado de acordo com o mercado e as decisões

Decisão	Lucro/Prejuízo C ₁ : aquecimento mercado	Lucro/Prejuízo C ₂ : mercado estável	Valor Monetário Esperado – VME
Importar	US\$300.000,00	US\$246.000,00	US\$289.200,00
Não importar	US\$18.000,00	US\$18.000,00	US\$18.000,00
Probabilidade	80%	20%	—

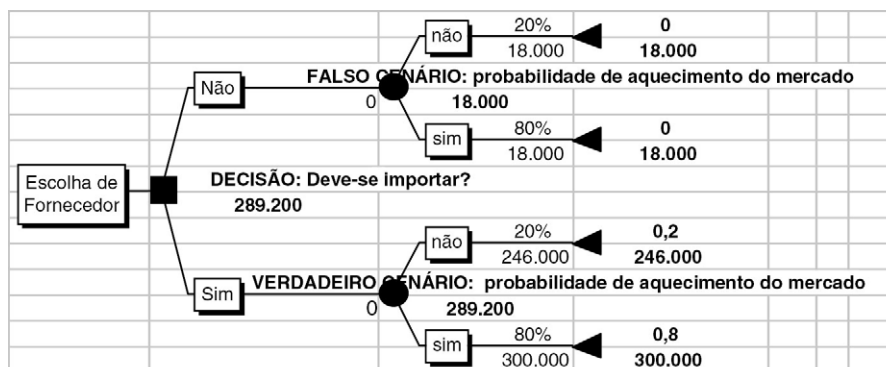


FIGURA 2.2 Primeira etapa da construção da árvore de decisão.

Utilizando a tabela de *payoff*, tem-se, na Figura 2.2, o desenvolvimento da Árvore de Decisão baseando-se nos dados da primeira decisão.

Observa-se, na Figura 2.2, que o ramo da decisão que leva o nome “VERDADEIRO” deve ser considerado como a decisão correta. Para finalizar a construção da Árvore de Decisão, incorporando os cenários de conjuntura do fornecedor estrangeiro, é necessário fazer a análise de atributos e custos envolvidos na avaliação de uma proposta de fornecimento, para explicitar a margem de contribuição de cada fornecedor e, subsequentemente, atribuir a probabilidade de risco conjuntural.

Passo 6: Levantamento dos atributos no processo de escolha do fornecedor

Conforme os procedimentos descritos, identificou-se, primeiramente, o problema gerado pela situação presente na empresa importadora. Para atender a demanda crescente do mercado, a empresa distribuidora de aço para construção visa aumentar seus estoques.

O fato de não conseguir o produto junto aos grandes fabricantes nacionais, que o usam para consumo próprio, ou exportam a maior parte de sua produção, leva a empresa a buscar alternativas de fornecimento. Segundo o levantamento de fabricantes de aço mundiais, foram selecionados os seguintes países que possuem fábricas que possam atender a demanda do mercado brasileiro: França, China e Ucrânia. Os critérios importantes na avaliação de um fornecedor são citados e analisados como atributos do produto no Quadro 2.1.

A decisão a ser tomada consiste na escolha de um dos três fornecedores fabricantes de aço para construção – francês, chinês ou ucraniano –, sendo que a indústria chinesa é conhecida por sua produção crescente e seus preços acessíveis; a indústria ucraniana é conhecida mundialmente por sua qualidade e seus preços competitivos; e a indústria francesa, preferida por sua qualidade e tradição.

A simplificação de nomear fornecedores de acordo com um país que identifique só a sua origem deve-se ao fato de que os preços de vários fornecedores do mesmo país tendem a um valor similar. Entretanto, a média de preços de um país difere da média do outro.

QUADRO 2.1 Critérios de avaliação do fornecedor

Critérios	Opção não desejável	Opção intermediária	Opção excelente
Preço comparado com <i>target price</i> ⁹ , baseado no preço do fornecedor nacional	Acima do <i>target price</i>	Igual ao <i>target Price</i>	Abaixo do <i>target price</i>
Qualidade de acabamento final do produto	Ruim	Boa	Excelente
Qualidade de acordo com exigência do mercado brasileiro	Não corresponde	Corresponde parcialmente	Corresponde totalmente
Prazo de entrega	Próximo semestre	Entrega no prazo de 30 dias	Entrega imediata
Forma de pagamento	Pagamento antecipado	Carta de crédito ¹⁰	Pagamento a prazo
Embalagem do produto	Ausência da embalagem própria para transporte	Embalagem normal de transporte	Embalagem própria para exportação

Para o estudo de caso em questão, é mais viável fazer a análise de três países fornecedores do que expor os dados de todos os fornecedores já analisados por país. O levantamento dos atributos é feito para avaliação de cada fornecedor e pesará na hora de decisão sobre a melhor opção. O **Quadro 2.2** trata do enquadramento dos fornecedores de acordo com os critérios de avaliação.

Os três fornecedores escolhidos apresentaram propostas que se encaixam nas opções intermediária e excelente, pré-qualificando os fornecedores para o respectivo fornecimento. Os preços apresentados por todos os fornecedores estrangeiros são descritos de forma detalhada a seguir, estando abaixo daquele pago ao fornecedor nacional. Disso decorre que, descontando-se os custos envolvidos, será possível obter-se uma margem de contribuição maior do que com produto nacional, que é de aproximadamente 16,6%. A estimativa e análise dos custos e da margem de contribuição na importação estão apresentadas a seguir.

Quantidade a ser importada:	3 mil toneladas
Valor FOB ¹¹ da Importação (FOB):	Preço Unitário * Quantidade
Custo da Carta de Crédito (C/C):	3% do valor FOB

9. Preço desejado pelo comprador.

10. Documento de garantia bancária de pagamento para o exportador desde que sejam cumpridas as condições estabelecidas pelo importador; uma forma de pagamento que apresenta segurança para ambas as partes.

11. *Free On Board*, modalidade de venda da mercadoria cujo preço na fatura comercial do exportador incluirá os custos de colocação dentro do navio (transporte utilizado no estudo). Os custos que ocorrerem depois disso (fretes, transbordos, taxas, impostos e outros) são por conta do importador.

QUADRO 2.2 Enquadramento dos fornecedores a partir dos critérios de avaliação

Critérios	Opção não desejável	Opção intermediária	Opção excelente
Preço			Os três
Qualidade de acabamento final do produto			Os três
Qualidade de acordo com exigência do mercado brasileiro			Os três
Prazo de entrega		Chinês / Ucraniano	Francês
Forma de pagamento		Os três	
Embalagem do produto		Chinês	Francês / Ucraniano

Custos da Importação e outros custos (CI):	47% do valor FOB
Custo das Mercadorias Vendidas (CMV):	(FOB + C/C + CI)
Receita Líquida de vendas no mercado nacional (RLV):	US\$600,00/t * 3.000t = US\$1.800.000,00
Margem de contribuição na importação (MCI):	RLV – CMV
Margem de contribuição na importação (MCI) %:	MC/RLV
Margem de contribuição <i>a priori</i> (MC):	US\$300.000,00 (ver Tabela 2. 3)
Diferença entre MC <i>a priori</i> e MCI:	MC – MCI
Diferença entre MC <i>a priori</i> e MCI, %:	(MC – MCI)/RLV

A Tabela 2.5 mostra o resumo dos cálculos dos custos da importação e da margem de contribuição a ser auferida na operação de importação com cada um dos fornecedores.

Para a conclusão da análise, faz-se necessária a comparação da margem de contribuição do produto nacional com a do produto importado. *A priori*, sabe-se que a margem de contribuição da venda do produto do fornecedor nacional é de 16,6%, o que resulta no valor de US\$300.000,00 (ver Tabela 2.3). Já na Tabela 2.6 evidencia-se o cálculo da margem de contribuição do produto importado, bem como sua comparação com o produto nacional.

TABELA 2.5 Cálculo da margem de contribuição na importação

Fornecedor	Preço Unitário	FOB	C/C	CI	CMV	RLV	MCI
Francês	310	930.000	27.900	437.100	1.395.000	1.800.000	405.000
Chinês	295	885.000	26.550	415.950	1.327.500	1.800.000	472.500
Ucraniano	300	900.000	27.000	423.000	1.350.000	1.800.000	450.000

TABELA 2.6 Comparativo da margem de contribuição do produto nacional e do importado

Fornecedor	RIV	CMV	MCI	MCI (%)	MC	MC– MCI	MC–MCI (%)
Francês	1.800.000	1.395.000	405.000	22,5%	300.000	105.000	5,9%
Chinês	1.800.000	1.327.500	472.500	26,3%	300.000	172.500	9,7%
Ucraniano	1.800.000	1.350.000	450.000	25,0%	300.000	150.000	8,4%

Observa-se, na [Tabela 2.6](#), que a rentabilidade do produto importado é maior. Ao se trabalhar com o produto dos fornecedores francês, chinês e ucraniano, obtém-se uma margem de contribuição aproximadamente 6%, 10% e 8% maior do que a do produto nacional, respectivamente.

A decisão em favor do fornecedor chinês fica evidente, já que a margem de contribuição apresentada é a maior de todas. Porém, lembrando-se que se trata de uma decisão em condições de risco, deve-se avaliar a probabilidade de sucesso de cada fornecedor. Conforme visto na [Tabela 2.2](#), as probabilidades de sucesso/insucesso para cada fornecedor, em função do risco conjuntural de intervenção política, foram de 90% a 10% para o fornecedor francês, de 40% a 60% para o fornecedor chinês, e de 90% a 10% para o fornecedor ucraniano.

Havendo insucesso no fornecimento da mercadoria, o comprador incorre em custo de abertura da carta de crédito. A [Tabela 2.7](#) mostra o último elemento dessa problemática, o modelo monetário, que transforma a decisão em valores monetários esperados (VME).

Nota-se que o maior VME de US\$402.300,00 indica que se deve optar pelo fornecedor ucraniano. A última etapa do processo decisório é preencher a Árvore de Decisão, observando-se que o campo que representa o valor monetário de sucesso da decisão de importação já está preenchido com a margem de contribuição estabelecida, *a priori*, como valor igual a US\$300.000,00 (ver valores na [Tabela 2.4](#)).

TABELA 2.7 Payoff para decisão de escolha do fornecedor estrangeiro

Decisão de escolha do fornecedor	Probabilidade de sucesso	Lucro/prejuízo em cenário de sucesso (US\$)	Probabilidade de insucesso	Lucro/prejuízo em cenário de insucesso (US\$)	VME (US\$)
Francês	90%	405.000	10%	–27.900	361.710
Chinês	40%	472.500	60%	–26.550	173.070
Ucraniano	90%	450.000	10%	–27.000	402.300

O cuidado a ser tomado é não incorrer em duplicidade de valores. Por isso, deve-se anular o valor de margem de contribuição no cenário de mercado (US\$ 300.000,00 e US\$246.000,00) e lançar o valor da margem de contribuição no campo correspondente de cada fornecedor. Desta maneira, a probabilidade de ocorrência de sucesso do fornecedor com melhor oferta será a seguinte:

- Probabilidade de aquecimento do mercado de $0,8 * \text{Probabilidade de sucesso de } 0,9 = 0,72$.
- Probabilidades de sucesso dos outros fornecedores são calculadas da mesma maneira e o mesmo procedimento vale para o ramo de combinação decisão – cenário “Importar – Mercado estável”.

A margem de contribuição nessa combinação terá descontado a oportunidade de remuneração do capital perdida (ver Tabela 2.3), conforme se encontra mensurado na Tabela 2.8.

Nota-se que os valores mensurados na Tabela 2.8 vão gerar a melhor alternativa em decorrência de o cenário de mercado permanecer estável. Preenchendo os dados na árvore, tem-se o caminho para a melhor decisão a ser tomada, tal como mostra a Figura 2.3.

A ferramenta apresenta os resultados obtidos na tabela *payoff* do fornecedor (ver a Tabela 2.7) e o ramo “VERDADEIRO” indica a melhor alternativa, com VME de US\$402.300,00, do fornecedor ucraniano no cenário de aquecimento do mercado.

A análise da Figura 2.3 apresenta os resultados para ambas as conjunturas, tanto para aquecimento do mercado como para permanência do mercado estável. Têm-se claros os resultados para alternativas de escolha de qualquer opção analisada. O método utilizado apresenta resultados monetários após revelação de informação sobre o mercado:

- Optando por importar e o mercado permanecendo estável – seguindo o ramo “VERDADEIRO”, o maior valor monetário apresentado é de US\$358.034.
- Optando por importar e ocorrendo o aquecimento do mercado – o ramo “VERDADEIRO” que apresenta o maior valor monetário é de US\$402.300.

TABELA 2.8 Margem de contribuição para decisão importar – cenário mercado estável

Decisão de escolha do fornecedor	Custo das mercadorias vendidas (US\$)	Margem de contribuição (US\$)	Oportunidade remuneração do capital por três meses, 1,2% a.m. (US\$)	Lucro/prejuízo do cenário de mercado estável (US\$)
Francês	1.395.000	405.000	50.825	354.175
Chinês	1.325.500	472.500	48.292	424.208
Ucraniano	1.350.000	450.000	49.185	400.815

A possibilidade de alterar valores e, principalmente, as probabilidades, à medida que são recebidas novas informações, proporciona flexibilidade gerencial ao tomador de decisão. Desta forma, o modelo proposto torna-se uma ferramenta ágil e precisa.

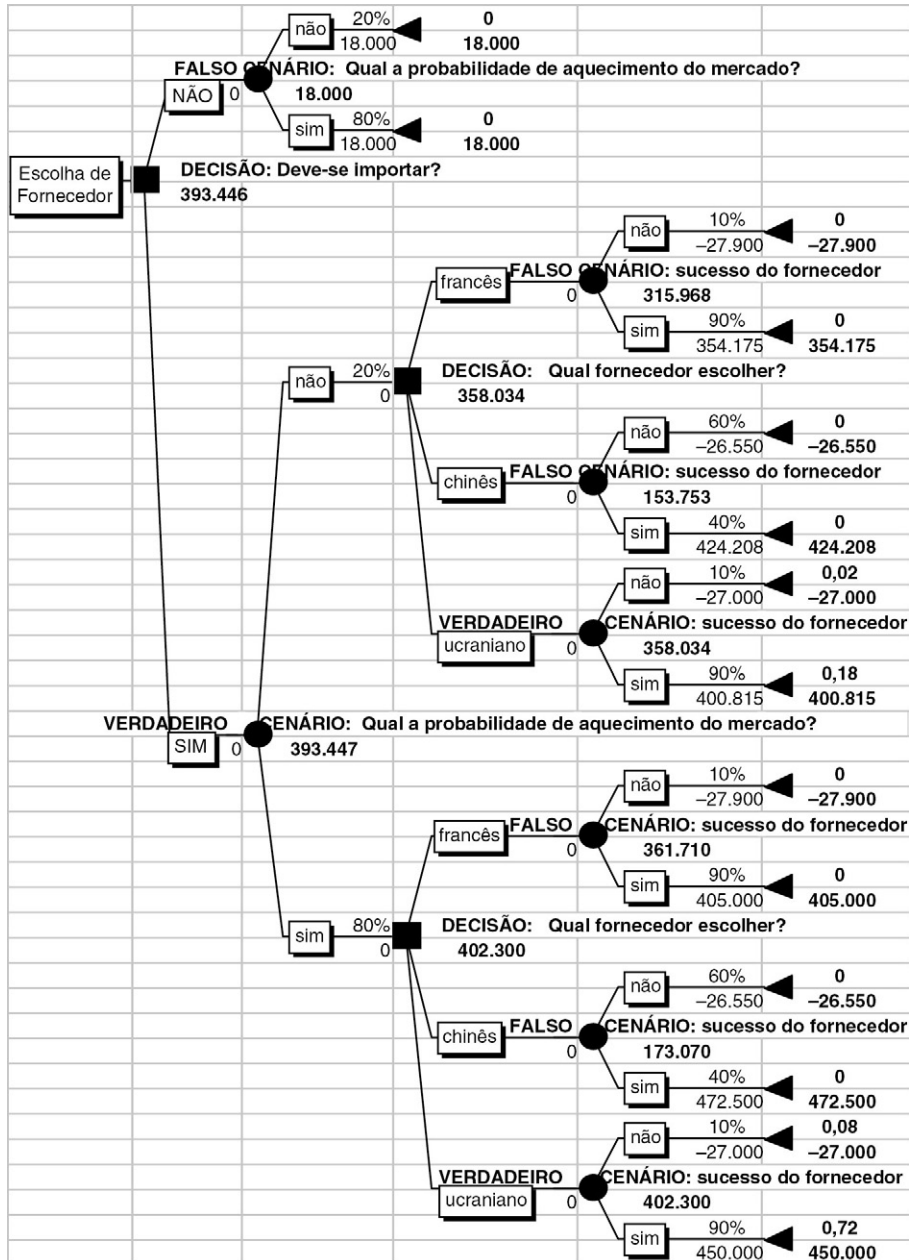


FIGURA 2.3 Estruturação completa da árvore de decisão para escolha de opções.

Vencida a principal etapa de modelagem do problema, considerada a mais complexa, o administrador poderá recorrer ao modelo já elaborado quantas vezes seja necessário.

Vale salientar que, devido à complexidade do desenvolvimento do modelo baseado no estudo de caso, algumas abordagens foram simplificadas. Dentre as principais limitações, destacam-se:

- o número de dados quantitativos sobre as ofertas de fornecedores estrangeiros, reduzido às informações essenciais, com o intuito de facilitar o entendimento do problema no nível macro;
- a avaliação do grau de incerteza por meio da atribuição de probabilidade, reduzida a dois cenários: de mercado e da conjuntura de política industrial no país de origem. Isso foi feito com a ressalva de que árvores muito complexas são de difícil análise e visualização.

CONSIDERAÇÕES FINAIS

Pode-se afirmar que o uso de modelos de decisão no processo decisório resulta em decisões eficazes, mesmo que o resultado, à primeira vista, não apresente a alternativa mais lucrativa. O modelo traduz a melhor opção em valores monetários esperados, considerando certas probabilidades de sucesso e insucesso. Portanto, a Teoria da Decisão deve ser utilizada nos processos decisórios em condições de risco para gerar respostas assertivas baseadas em métodos e evitar julgamentos puramente intuitivos dos tomadores de decisão.

Um dos impedimentos para a difusão do método seria a falta de habilidade junto à ferramenta por parte do tomador de decisão, que é adquirida com acompanhamento do especialista e depois de algum tempo de uso. Os conceitos de nó de decisão e de nó de cenário devem ser bem entendidos e diferenciados na aplicação do problema.

Em termos de julgamentos gerenciais no processo de decisão, os decisores não devem deixar de considerar a possibilidade da presença de vieses ou pré-conceitos, sob pena de afetarem e comprometerem o processo como um todo. Os decisores envolvidos no processo de tomada de decisão devem ter em mente que, na busca pela melhor decisão, devem trabalhar com todas as informações disponíveis, inclusive aquelas consideradas difíceis.

Não obstante os problemas de construção da Árvore de Decisão, seria injusto não mencionar as vantagens desse método, que consiste em um processo relativamente rápido de aprendizagem, pois a construção de regras acontece em linguagem de usuário, além de ser possível gerar regras em áreas em que um profissional tem dificuldades de estruturar seus conhecimentos. É conveniente para a construção de modelos não paramétricos e mesmo assim possui uma clara classificação e alto nível de predição.

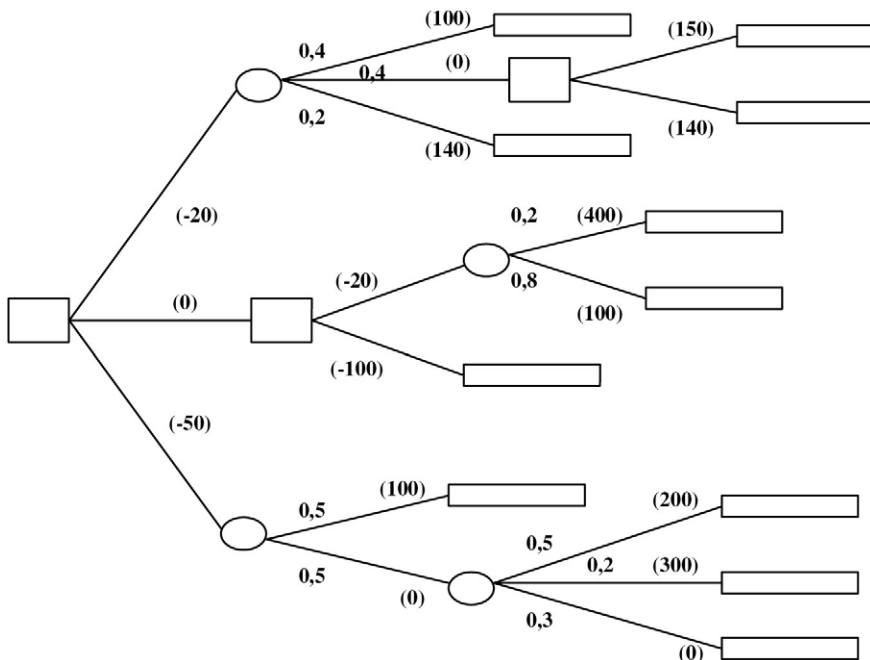
Nesse contexto, a Árvore de Decisão é uma ferramenta de apoio no processo de tomada de decisão. Muitos softwares destinados à análise de dados já dispõem de técnicas de construção de Árvores de Decisão. Em áreas em que o preço de um erro é muito alto, são de grande ajuda no trabalho do analista ou do gerente.

Cabe destacar que as Árvores de Decisão garantem maior flexibilidade sobre os mecanismos tradicionais à medida que as decisões são definidas com maior grau de flexibilidade, visto que muitos eventos podem ser mapeados ao longo dos diversos ramos, envolvendo várias decisões de sequenciamento. Além disso, a Árvore de Decisão acaba incorporando conceitos de incerteza na análise de alternativas de investimentos ou financiamentos, o que a torna uma técnica muito eficaz.

Uma grande vantagem da aplicação da ferramenta quantitativa apresentada neste capítulo refere-se ao fato de que uma decisão de priorização de determinado investimento em relação a outros não é apresentada de forma intuitiva em uma “caixa preta”. Ela se encontra amparada em uma modelagem matemática robusta e transparente, que apresenta um encadeamento lógico e racional de premissas e preferências por parte do decisor acerca de cada um dos atributos considerados durante o processo.

EXERCÍCIOS

1. Resolva, sem o auxílio de nenhum software, a árvore a seguir:



2. Novembro de 2012. Um estudante precisava decidir o que fazer com seu Corsa 2009. O gerente da oficina mecânica com que trabalha lhe telefonara pela manhã dizendo que, ao fazerem a revisão de rotina no carro, verificaram que o motor não estava em boas condições e, para evitar um dano permanente, era preciso

um serviço imediato que custaria, provavelmente, R\$ 1.200. Foi solicitada autorização para fazer o serviço, mas o estudante mostrou-se indeciso, pois sua situação financeira era crítica, pelo menos até quando voltasse a trabalhar, em novembro de 2013. Devido a este fato, ele resolveu minimizar suas despesas com o carro durante os 12 meses seguintes. Para poder tomar a melhor decisão, o estudante passou a manhã investigando as alternativas possíveis. Por volta do meio-dia, concluiu que deveria optar por ou trocar o seu Corsa por um Palio novo, ou manter o Corsa. Se decidisse pela troca, evitaria o conserto. O dispêndio de caixa líquido com o novo carro, dando o Corsa como troca, seria R\$ 3.000, que ele precisaria pagar imediatamente. Se decidisse manter o Corsa, ainda não estava convencido se deveria fazer o conserto de R\$ 1.200. Mesmo que autorizasse o conserto, continuaria tendo que fazer reparos anuais de cerca de R\$ 1.000 no carro. Se decidisse não consertá-lo, ele achava que havia uma chance de apenas 40% de o carro quebrar e uma chance de 60% de ele não quebrar. Se ocorresse um enguiço sério, parecia igualmente provável que o conserto custaria R\$ 2.000 ou R\$ 3.000. Como ambos os carros teriam valores de mercado mais ou menos iguais em um ano, o estudante achava que este fato não precisaria ser considerado. No entanto, ele achava que os custos com combustível eram relevantes. Estimou-os em R\$ 2.400 para o Corsa e em R\$ 1.100 para o Palio, no decorrer dos 12 meses seguintes. Analise o problema de decisão e recomende uma estratégia a ser seguida.

3. Um explorador de petróleo precisa decidir se perfura ou não determinado local. Ele não está certo se a perfuração será seca ou molhada. Seus lucros são determinados pelas suas ações e pela natureza do poço: o custo de perfuração é de R\$ 700.000; se um poço molhado for perfurado (50% de probabilidade), o lucro bruto é de R\$ 1.700.000.
 - a. O poço deve ou não ser perfurado?
 - b. Por um custo de R\$ 100.000, o explorador poderia fazer uma série de sondagens sísmicas capazes de determinar a estrutura geológica do local. As sondagens mostrarão se o terreno subterrâneo tem: (i) nenhuma estrutura (ruim) – 40% de probabilidade; (ii) estrutura aberta (mais ou menos) – 35% de probabilidade; ou (iii) estrutura fechada (muito bom) – 25% de probabilidade. A probabilidade de o poço estar seco depende do resultado do teste: (i) teste sísmico revelando “nenhuma estrutura” – 75%; (ii) teste sísmico revelando “estrutura aberta” – 50%; ou (iii) teste sísmico revelando “estrutura fechada” – 10%. O teste sísmico deve ser feito? Qual é a ação ótima agora?
 - c. Se o teste sísmico fosse perfeito, ou seja, indicando sempre “nenhuma estrutura” para um poço seco e “estrutura fechada” para um poço molhado, até que valor estaria o explorador disposto a pagar por este teste?
4. Um corretor de imóveis recebe uma proposta para colocar três salas comerciais à venda. Se ele aceitar o negócio, terá que desembolsar R\$ 700 para aquisição de um terno e realização de um pequeno reparo no seu automóvel. Além disso, o custo para tentar vender cada imóvel (ligações telefônicas, deslocamento, tempo

despendido etc.) pode ser estimado como equivalente a 0,40% do preço de venda do imóvel. O corretor é obrigado a começar as tentativas de venda pelo imóvel A e nunca deve tentar vender dois imóveis simultaneamente. Se e quando ele vender o imóvel A, poderá escolher qual imóvel tentará vender em seguida. Mas se, após um mês, falhar na tentativa de vender o imóvel A (ou qualquer outro), não poderá continuar tentando vender os outros imóveis. Mas ele também pode desistir de continuar tentando quando quiser; não precisa esperar um fracasso. Sua comissão de venda é de 4% sobre o valor do imóvel e os demais dados relevantes encontram-se a seguir.

Imóvel	Preço	Comissão	Custo	Pr (sucesso)
A	R\$ 25.000,00	R\$ 1.000,00	R\$ 800,00	70%
B	R\$ 50.000,00	R\$ 2.000,00	R\$ 200,00	60%
C	R\$ 100.000,00	R\$ 4.000,00	R\$ 400,00	50%

- Ele deve aceitar o negócio? Se sim, em que ordem deve tentar vender os outros imóveis? Em que momento deve parar?
- O corretor está inseguro quanto à probabilidade de vender o imóvel A. Para quanto este parâmetro deveria mudar para mudar a decisão do corretor de aceitar ou não o negócio? Quanto passaria a valer o negócio para o corretor para diferentes valores deste parâmetro?
- Faça uma análise de sensibilidade completa em relação à probabilidade de venda dos três imóveis. Considere que a probabilidade de vender A é, no mínimo, 30% e, no máximo, 90%; a de vender B é, no mínimo, 30% e, no máximo, 80%; e a de vender C é, no mínimo, 25% e, no máximo, 75%. Qual dos três parâmetros apresenta maior potencial para impactar o valor do negócio?
- Agora considere o perfil de aversão a risco do corretor: para ele, ter 50% de chance de ganhar R\$ 1.600 e 50% de chance de ganhar R\$ 5.600 não equivale a ganhar, com certeza, R\$ 3.600 (o valor esperado), mas, sim, R\$ 2.500. Responda novamente as questões levantadas na letra a.

REFERÊNCIAS

- ALBRIGHT, S.; WINSTON, W.; ZAPPE, C. *Data analysis & decision making with Microsoft Excel*. 2. ed. [SI]: Thomson Brooks/Cole, 2003.
- CARVALHO, I. *Métodos de mineração de dados (Data Mining) como suporte à tomada de decisão*. Dissertação de Mestrado, ITA, São José dos Campos, SP, 2002.
- CORRAR, L.; THEÓPHILO, C. *Pesquisa operacional para decisão em contabilidade e administração*. São Paulo: Atlas. 2004.
- GALESNE, A.; FENSTERSEIFER, J.; LAMB, R. *Decisões de investimentos da empresa*. 2. ed. São Paulo: Atlas. 1999.
- GAMA, J. Árvore de decisão, 2000. Disponível em: <http://www.liacc.up.pt/~jgama/Mestrado/ECD1/Arvores.html>. Acesso em 14/08/2004.
- GOMES, L.; ARAYA, M.; CARIGNANO, C. *Tomada de decisões em cenários complexos*. São Paulo: Pioneira Thompson Learning. 2004.
- GOMES, L.; GOMES, C.; ALMEIDA, A. *Tomada de decisão gerencial*. São Paulo: Atlas. 2002.
- MINARDI, A. *Teoria de opções aplicada a projetos de investimentos*. São Paulo: Atlas. 2004.

PINDYCK, R.; RUBINFELD, D. *Microeconomia*. 2. ed. São Paulo: Makron Books. 1994.

PRADO, E.; YU, A. Análise de decisão na terceirização da tecnologia de informação: um estudo de caso no setor químico brasileiro. In: SEMINÁRIOS DE ADMINISTRAÇÃO DA USP – SEMEAD, 6, 2003, São Paulo. *Anais...* São Paulo: USP/FEA. 2003.

SECURATO, J. *Decisões financeiras em condições de risco*. São Paulo: Atlas. 1996.

TONETTO, L. M.; KALIL, L.; MELO, W.; SCHNEIDER, D.; STEIN, L. O papel das heurísticas no julgamento e na tomada de decisão sob incerteza. *Estudos de Psicologia*, v. 23, nº 2, 181-189, abr./jun. 2006.

KAHNEMAN, D.; TVERSKY, A.; Prospect theory: An analysis of decision under risk. *Econometrica*, v. 47, March 1979.

Simulação do escoamento de arroz para verificar se a intermodalidade compensa

Luís Otávio Marins Ribeiro

MADE/UNESA

Marco Aurélio Carino Bouzada

MADE/UNESA

Objetivos de aprendizagem do capítulo

- Saber modelar um problema real complexo em situação de risco através da Simulação de Monte Carlo.
- Aprender a gerar números aleatórios e a transformá-los em variáveis aleatórias com suas próprias distribuições de probabilidades.
- Entender como replicar a simulação e coletar e analisar as estatísticas de saída.
- Saber conduzir uma Análise de Sensibilidade acerca dos parâmetros mais importantes.
- Entender as vantagens de uma abordagem experimental em comparação a uma abordagem analítica no tratamento de problemas complexos.

INTRODUÇÃO

A logística vem mostrando que seu uso adequado dentro das empresas pode trazer benefícios aos processos, com consequente redução nos custos. O processo mais custoso para as empresas é o de transporte. A atividade de transporte representa até 60% dos custos logísticos.¹ Não só no Brasil, mas em todo o mundo, busca-se a melhoria na infraestrutura e nos processos dos transportes, na tentativa de redução nos custos.

Já ouvimos há muito tempo que o escoamento da produção no Brasil não funciona adequadamente, seja pelo estado de má conservação das estradas, seja por não existir outro modal, com a mesma disponibilidade do rodoviário, que possa escoar esta produção.

Como a malha ferroviária brasileira está concentrada nas regiões Sul e Sudeste, seria necessária a utilização de outros modais que permitissem acessar as demais

1. Gomes; Ribeiro (2004).

TABELA 3.1 Comparação das principais características dos modais²

	Ferroviário	Rodoviário	Marítimo	Dutoviário	Aéreo
Velocidade	3°	2°	4°	5°	1°
Disponibilidade	2°	1°	4°	5°	3°
Dependência	3°	2°	4°	1°	5°
Capacidade	2°	3°	1°	5°	4°
Frequência	4°	2°	5°	1°	3°
Custo fixo	4°	1°	2°	5°	3°
Custo variável	3°	4°	2°	1°	5°

² Adaptação de Bowersox; Closs (2001).

regiões. Estes modais poderiam ser o rodoviário e o hidroviário, por exemplo. Na verdade, por que o transporte de carga precisa ficar restrito a uma modalidade apenas de transporte se cada uma delas apresenta diferentes vantagens, tornando-se mais adequada, portanto, a trechos, circunstâncias e situações específicas?

Para ilustrar essas diferenças de adequação entre os modais, é apresentada, a seguir, a [Tabela 3.1](#), que compara as características mais relevantes dos modais de transportes, dando, assim, uma visão das diferenças entre eles, possibilitando uma melhor decisão em relação a que modal utilizar e em que momento.

Partindo desse ponto, surge a pergunta: seria possível reduzir os custos com o uso da intermodalidade de transporte?³ Ou, ainda, seria possível reduzir os custos logísticos de transportes, para uma grande quantidade de carga e em um longo percurso, com o uso da intermodalidade de transporte, sem impactar de forma significativamente prejudicial a qualidade do nível de serviço?

Mas quais seriam as possibilidades adicionais de modais a serem utilizados? Dos cinco modais utilizados para o transporte de carga, deu-se ênfase ao rodoviário, ao ferroviário e ao hidroviário. Os outros dois ficaram de fora, pois o aéreo apresenta custo demasiadamente alto e o por dutos é por demais limitado a produtos bem específicos.

Dessa forma, a sugestão consiste em dividir os percursos mais longos entre os modais rodoviário, ferroviário e hidroviário (não necessariamente nessa ordem) e, assim, utilizar as vantagens de cada um para compensar as desvantagens apresentadas pelo outro. Ou seja, se nos detivermos ao que há de melhor em cada modal, poderemos utilizar os pontos fortes de cada um, sobrepujando os pontos fracos de cada modal utilizado.

3. Integração dos serviços de mais de um modo de transporte, desde o momento da carga inicial até a descarga final.

Para ajudar a responder às perguntas colocadas anteriormente, propomos a utilização da ferramenta de Simulação de Monte Carlo, uma vez que esta metodologia já é utilizada em estudos logísticos com bastante frequência.

SIMULAÇÃO DE MONTE CARLO

A Simulação de Monte Carlo é um método de avaliação interativa de um modelo determinístico que usa números randomizados como entradas. Esse método é mais utilizado quando o modelo é complexo, ou não linear, ou quando envolve um número razoável de parâmetros de incerteza. Uma simulação pode envolver mais de 10 mil avaliações do modelo estudado, uma tarefa difícil que no passado só poderia ser realizada por supercomputadores.⁴

O método de Monte Carlo, um entre vários voltados para análise da propagação da incerteza, tem como grande vantagem poder determinar como uma variação randomizada, já conhecida, ou o erro afetam o desempenho ou a viabilidade do sistema que está sendo modelado.⁵

Atualmente, a complexidade dos processos logísticos faz com que as ferramentas estáticas e determinísticas tenham aplicabilidade pontual. A Simulação é uma ferramenta que leva em consideração a dinâmica e a aleatoriedade presentes nos processos logísticos, avaliando-os com precisão muito maior e permitindo ganhos substanciais nas operações.⁶

A Simulação é a técnica mais utilizada no planejamento de sistemas logísticos pelo fato de poder abordar questões referentes à escolha dos **modais de transporte** e da localização de armazéns, aos custos de transportes e de imobilização, ao nível de serviço aos clientes etc. Torna-se possível, assim, avaliar os diversos componentes do sistema e optar pelas formas mais adequadas de operação.⁷

A Simulação tem acompanhado a evolução da importância da Logística no meio empresarial. Necessidades de alocação de espaço, requisitos de movimentação cada vez mais variados são apenas algumas das questões de distribuição e armazenagem que podem ser tratados. Uma vez que a Logística, em muitos aspectos, está relacionada à movimentação de operações envolvendo o tempo, às mudanças de locais e aos eventos probabilísticos, a modelagem dinâmica se torna fundamental. A principal vantagem da aplicação da Simulação em sistemas logísticos consiste no fato de permitirem testar virtualmente as alternativas de operação.⁸

Podemos entender, então, a Simulação como uma das ferramentas mais apreciadas para os estudos de processos logísticos. Ela nos permite representar os problemas de forma mais próxima à realidade, em função da quantidade de dados utilizados. E os resultados, então, são avaliados com maior precisão. de forma mais precisa.

4. Shamblin; Stevens Jr. (1979).

5. Moore; Weatherford (2005).

6. Chwif (2008).

7. Johnson; Wood (1996).

8. Johnson; Wood (1996).

A SIMULAÇÃO DO ESCOAMENTO DA PRODUÇÃO DE ARROZ

Comparada a outras ferramentas, a Simulação demonstra boa representação das situações complexas e incertas do mundo real. Por isso, é considerada de grande importância neste estudo para ajudar na verificação, entre outros impactos, da possível redução de custos logísticos a partir da utilização da intermodalidade de transporte.

Assim, foi escolhida para este estudo a técnica de Simulação como o mecanismo metodológico, em função da sua capacidade para proporcionar a visão sistêmica sobre a realidade operacional. A ideia consiste em estabelecer uma metodologia capaz de obter uma confirmação referente à possibilidade de redução dos custos logísticos através do uso da intermodalidade de transporte. Apesar de procurar avaliar também outros indicadores de desempenho, a ênfase está na análise dos custos de transportes, de modo a auxiliar a identificação da melhor forma de agrupar os modais de transportes no escoamento da produção, podendo assim alcançar a otimização na distribuição dos produtos.

A Simulação foi aplicada a partir de parâmetros reais e ambientada em uma planilha do Excel, ilustrada parcialmente na [Figura 3.1](#). Estamos considerando o transporte da produção de arroz, partindo do Vale do Jacuí (RS) até a Região dos Lagos (RJ).

Na [Figura 3.1](#) são apresentadas as variáveis aleatórias utilizadas para aplicação da ferramenta de Simulação no caso do transporte unimodal⁹ rodoviário.

A carga a ser transportada é de 1.200t, por um trecho de 1.922km, percurso a ser percorrido em dois dias, de acordo com a exigência do cliente. A velocidade média do caminhão foi considerada como uma variável aleatória, com Distribuição Triangular, de mínimo de 40km/h, máximo de 70km/h e moda de 60km/h. Além da velocidade média, foram consideradas como variáveis aleatórias (também com Distribuição Triangular): as horas úteis por dia, sendo o mínimo de 12 horas, máximo de 17 horas e moda (valor mais provável) de 15 horas; e o custo de transporte por tonelada, sendo este modelado com um mínimo de R\$ 55, máximo de R\$ 78 e moda de R\$ 60.

Foram simuladas mil ocorrências para que o experimento pudesse ser validado estatisticamente. De cada uma delas sorteou-se a velocidade média e as horas úteis por dia, tendo sido calculada a partir delas a duração do trajeto em horas e o consequente nível de serviço. Para efeitos deste estudo, foi considerado apenas o aspecto de nível de serviço referente ao percentual de vezes em que o prazo estipulado pelo cliente – dois dias – foi cumprido. Também foi sorteado o custo por tonelada e, a partir dele, calculado o custo total.

Em seguida, foi abordado o caso da intermodalidade, em que os parâmetros também foram baseados em dados reais. Para este caso, também foi utilizada a Distribuição Triangular para modelar o comportamento de todas as variáveis aleatórias.

O primeiro trecho foi considerado como realizado pelo modal rodoviário, e a distância do percurso que vai de Cachoeira do Sul até Porto Alegre foi levantada, respeitando os outros parâmetros utilizados pelo modal em questão na primeira simulação. O trecho realizado pelo modal ferroviário foi de Porto Alegre (RS) até o Porto de São Francisco do Sul (SC). Por último, para o modal hidroviário foi

9. Utilizando apenas um modo de transporte.

Prazo	2 dias		Horas úteis por dia							
			Mínimo	12 horas						
Carga	1.200 ton		Máximo	17 horas						
Distância	1.922 km		Moda	15 horas						
Velocidade média			Custo/ton							
Mínimo	40		Mínimo	R\$ 55,00						
Máximo	70		Máximo	R\$ 78,00						
Moda	60		Moda	R\$ 60,00						
Rodoviário										
Simulação	Aleatório	Velocidade	Duração (h)	Aleatório	Horas úteis por dia	Duração (d)	Chegou?	Aleatório	Custo/ton	Custo Total
988	0,97	66,82	28,77	0,44	14,56	1,98	1	0,48	R\$ 62,71	R\$ 75.251,31
989	0,99	67,96	28,28	0,31	14,15	2,00	1	0,74	R\$ 61,08	R\$ 73.297,22
990	0,29	53,12	36,19	0,95	16,33	2,22	0	0,24	R\$ 73,66	R\$ 88.395,31
991	0,75	61,31	31,35	0,18	13,62	2,30	0	0,21	R\$ 59,50	R\$ 71.398,74
992	0,99	68,12	28,21	0,61	15,04	1,88	1	0,09	R\$ 63,40	R\$ 76.085,28
993	0,05	45,25	42,47	0,15	13,50	3,15	0	0,21	R\$ 59,15	R\$ 70.983,56
994	0,27	52,76	36,43	0,01	12,30	2,96	0	0,36	R\$ 57,72	R\$ 69.258,20
995	0,74	61,19	31,41	0,35	14,31	2,20	0	0,47	R\$ 61,65	R\$ 73.981,70
996	0,88	63,90	30,08	0,47	14,65	2,05	0	0,20	R\$ 62,33	R\$ 74.792,02
997	0,33	53,97	35,61	0,44	14,57	2,44	0	0,04	R\$ 62,12	R\$ 74.538,70
998	0,66	59,97	32,05	0,68	15,20	2,11	0	0,80	R\$ 66,43	R\$ 79.713,64
999	0,36	54,67	35,16	0,43	14,53	2,42	0	0,47	R\$ 62,59	R\$ 75.105,07
1.000	0,26	52,45	36,64	0,79	15,56	2,35	0	0,12	R\$ 64,55	R\$ 77.461,73
Média	0,50	56,59	34,39	0,48	14,59	2,37	10%	0,51	R\$ 63,52	R\$ 76.222,93

FIGURA 3.1 Ilustração parcial da planilha de simulação (alternativa unimodal).

considerado o percurso do Porto de São Francisco do Sul (SC) até o Porto do Forno em Arraial do Cabo (RJ). Ainda foram levados em consideração os custos de transbordo, do caminhão para o trem e do trem para o navio, considerando as variáveis mencionadas anteriormente.

Torna-se relevante ressaltar que foi considerado o uso do contêiner para o transporte da carga, uma vez que a sua utilização facilita o transbordo.

Também no caso da intermodalidade, foram simuladas mil ocorrências do experimento. Para o modal rodoviário, utilizou-se o mesmo procedimento do caso unimodal. Para cada um dos outros modais, o ferroviário e o hidroviário, sorteou-se apenas a velocidade média e o custo por quilômetro por cada mil toneladas, uma vez que consideramos úteis as 24 horas do dia. Em seguida, foi calculada a duração da viagem.

No caso do transbordo para ambos, houve o sorteio da duração unitária (em minutos por tonelada) e, a partir dela, o cálculo da duração total em dias. O custo de cada transbordo foi calculado a partir da respectiva duração e do custo unitário (em reais por minuto).

Por fim, foram somadas as durações (em dias) das viagens empreendidas pelos três modais envolvidos na intermodalidade e dos transbordos. A partir deste total, foi gerado o nível de serviço intermodal (percentual de vezes em que o prazo estipulado pelo cliente – dois dias – foi cumprido). Também foram somados todos os custos envolvidos na simulação intermodal, incluindo transportes e transbordos.

Os principais indicadores de desempenho considerados na análise – duração média da viagem (em dias), nível de serviço e custo médio – foram coletados para o conjunto de mil simulações; os resultados estão apresentados na Tabela 3.2 a seguir.

Os resultados das alternativas unimodal e intermodal estão apresentados lado a lado (para comparação), além da diferença percentual ocasionada pelo uso da intermodalidade em cada indicador de desempenho.

Esses resultados, na verdade, são dinâmicos e mudam a cada vez que os valores da planilha são atualizados. O conteúdo da Tabela 3.2 é apenas uma fotografia, não captando, naturalmente, o aspecto dinâmico desses valores.

Como pode ser observado, os valores obtidos para cada um dos modos de transportes – intermodalidade (rodoviário, ferroviário e hidroviário) e unimodalidade (rodoviário) – são razoavelmente diferentes. No caso da duração média em dias, o intermodal apresenta 2,12 dias, ao passo que o unimodal apresenta 2,36 dias. Isto surpreende, pois o intermodal se mostra com um tempo médio de entrega menor; enquanto esperava-se que o unimodal o superasse neste indicador. No caso do nível de serviço, como consequência da duração média, a intermodalidade também supera a unimodalidade: temos 29,20% e 11,60%, respectivamente. Já em relação ao custo médio, era esperado um custo bem menor para a intermodalidade em relação ao unimodal e, o que podemos observar, embora seja um custo menor e significativo, a diferença não é tão grande assim, sendo de R\$ 63.910,78 para o intermodal e de R\$ 76.473,26 para o unimodal.

Após a simulação do cenário básico – carga de 1.200t e dois dias de prazo de entrega – foi feita uma análise de sensibilidade em relação a esses dois parâmetros – variando a carga para 600t e 2.400t e o prazo de entrega para três dias.

TABELA 3.2 Indicadores de desempenho para cada possibilidade de transporte (intermodal e unimodal) – utilizando dois dias como prazo de entrega e 1.200t para a carga

	Intermodal			Unimodal		Diferença (%)		
	Rodoviário	Transbordo	Ferroviário	Transbordo	Hidroviário	Total	Rodoviário	Diferença (%)
Duração média (d)	0,16	0,04	0,92	0,09	0,91	2,12	2,36	-10,18%
Nível de serviço						29,20%	11,60%	151,72%
Custo médio	R\$ 10.258,49	R\$ 87,81	R\$ 28.507,94	R\$ 401,90	R\$ 24.654,63	R\$ 63.910,78	R\$ 76.473,26	-16,43%

TABELA 3.3 Indicadores de desempenho para cada possibilidade de transporte intermodal e unimodal – análise de sensibilidade para o prazo de entrega e a carga transportada

		Duração média (d)	Nível de serviço	Custo médio
1.1.200t – 2 dias	Intermodal	2,12	29,20%	R\$ 63.910,78
	Unimodal	2,36	11,60%	R\$ 76.473,26
1.200t – 3 dias	Intermodal	2,13	100,00%	R\$ 63.890,69
	Unimodal	2,36	96,40%	R\$ 76.662,44
600t – 2 dias	Intermodal	2,08	37,20%	R\$ 32.074,10
	Unimodal	2,36	11,80%	R\$ 38.270,63
2.400t – 2 dias	Intermodal	2,23	7,30%	R\$ 127.873,30
	Unimodal	2,36	12,70%	R\$ 153.659,15

A simulação foi rodada novamente para cada cenário, e os resultados foram coletados, estando apresentados em conjunto com os do cenário básico na [Tabela 3.3](#).

É válido destacar o último cenário, em que a alternativa intermodal mostrou-se mais rápida (2,23 dias) – na média! – do que a alternativa unimodal (2,36), mas encerrou um nível de serviço inferior (7,30% contra 12,70%). Tal aparente paradoxo deve-se à maior variabilidade das durações das viagens unimodais, que apresentam valores em média um pouco mais altos, mas com ocorrências extremas, fazendo com que um percentual maior das viagens dure muito ou muito pouco, ficando abaixo da barreira dos dois dias mais vezes do que no caso intermodal.

Tal tipo de constatação só foi possível devido ao uso da Simulação, visto que, normalmente, através de abordagens analíticas só se consegue observar o comportamento médio das grandezas relevantes.

CONSIDERAÇÕES FINAIS

Os resultados deste estudo foram obtidos a partir da Simulação com dados reais e permitiram verificar a viabilidade da intermodalidade em uma situação real de escoamento da produção de arroz – de Cachoeira do Sul (RS) até Arraial do Cabo (RJ). Este era o principal objetivo da pesquisa – a análise a respeito das vantagens e das desvantagens oriundas do uso da intermodalidade, neste caso específico, além de validar a utilização da Simulação, instrumento que foi capaz de considerar as incertezas presentes nesta operação logística e de subsidiar com informações úteis para o processo decisório, como o impacto no custo e em outros indicadores de desempenho.

Em relação à incerteza, o exemplo real apresentado neste trabalho considerou estocástico o comportamento de diversas grandezas capazes de impactar o desempenho do sistema: velocidade média de transporte e de transbordo, horas disponíveis por

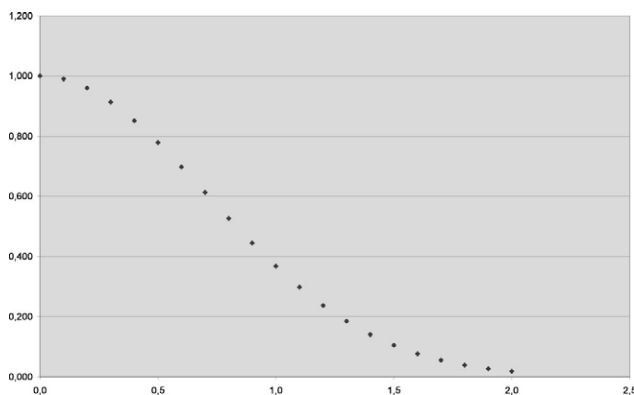
dia para o modal rodoviário e custos unitários de transporte e de transbordo. Tais grandezas foram modeladas a partir da Distribuição Triangular de probabilidades, mas a flexibilidade da Simulação permite que qualquer distribuição – principalmente nos casos em que nenhuma distribuição teórica represente bem o comportamento da grandeza – seja usada. Tal característica da ferramenta permitiu que a realidade fosse modelada de forma bem acurada, uma vez que houve um esforço na coleta dos dados necessários para uma representação realista do problema em questão.

A partir desses resultados, foi possível observar que, no caso da duração média, a alternativa intermodal apresenta uma quantidade em dias inferior à da unimodal. Esperava-se que, com o uso da intermodalidade, o nível de serviço caísse um pouco, sendo tal queda irrelevante diante da redução dos custos. Mas em relação a esse atributo, como consequência da duração média, a intermodalidade também supera a unimodalidade na maioria dos cenários, o que causou certa surpresa. Tal fato pode ser explicado em parte: (i) pela presença – no trecho em questão – de muitas montanhas no caminho dos caminhões, diminuindo um pouco a velocidade média; (ii) e pelo desenho recortado do litoral brasileiro e da sua malha viária costeira, que impossibilita os caminhões de percorrerem uma linha reta (o que não acontece com os navios), tornando a distância unimodal maior do que o somatório das distâncias dos trechos da alternativa intermodal.

Já em relação ao custo médio, esperava-se um custo menor para a intermodalidade em relação ao unimodal, o que foi possível se observar em todos os cenários. Cabe aqui uma ressalva, no sentido de que tais conclusões sejam consideradas com precaução. Esta pesquisa visou analisar apenas questões objetivas e quantitativas relacionadas à distribuição, como custo e duração da viagem. Mas é claro que a alternativa intermodal – apesar de ter se apresentado mais vantajosa nesses aspectos – envolve alguns inconvenientes potenciais – como aspectos humanos relacionados ao comodismo e aversão à complexidade, além de outros também menos tangíveis, envolvidos nos transbordos e em outros momentos da operação intermodal – que não foram considerados e que podem constituir parte da causa da restrita utilização da intermodalidade no mundo real, ao menos no Brasil.

EXERCÍCIOS

1. Existe interesse no cálculo da área sob a curva e^{-x^2} , entre os pontos 0 e 2.



Como não existe uma função predefinida para o cálculo da área de uma figura irregular como essa, precisaríamos obter a integral da função e calcular o seu valor entre os pontos 0 e 2. O problema é que tal função não pode ser integrada; em outras palavras, não é possível obter uma função analítica geral que represente a integral da função em questão. O que se sabe (e pode ser demonstrado) a respeito dessa função é que a sua integral definida – entre 0 e $+\infty$ – é igual à metade da raiz quadrada de π , ou 0,8862, aproximadamente. No entanto, não queremos saber a área desde 0 até o infinito, mas apenas até o ponto em que $x = 2$. Obviamente, este valor que buscamos é menor do que 0,8862, que pode ser considerado um limite superior para a área procurada.

- a. “Distorça” a realidade e dê à figura uma forma da qual sejamos capazes de calcular a área: o triângulo. Para tal, basta “fingirmos” que a curva é uma reta e, assim, teremos um triângulo formado por tal segmento de reta e os dois eixos. Calcule a área desse triângulo.
 - b. Agora descubra a área procurada de maneira experimental. Simule o sorteio de 10 mil pontos dentro do retângulo imaginário $(0;0) - (2;1)$, que engloba a área procurada. Verifique que percentual (desses 10 mil pontos) caiu abaixo da curva (ou seja, dentro da área visada), para ter uma boa ideia da área que procuramos, já que a área do retângulo é conhecida.
 - c. Compare as duas abordagens.
2. O proprietário de uma empresa que vende material de camping está empenhado em melhorar o atendimento da demanda de barracas, produto que tem importante participação na receita da empresa. A margem de contribuição bruta unitária do produto é de R\$ 25,00. As perdas de vendas por falta de estoques têm sido constantes, já que, em geral, os clientes que procuram o produto querem levá-lo no mesmo dia e dificilmente aceitam esperar a chegada de novos pedidos. A capacidade do galpão (alugado) utilizado para estocagem das barracas é suficiente, mas R\$ 2,00 são cobrados por dia por barraca armazenada no início do período (ou seja, você deve considerar – para efeitos de cálculo do custo de estoque – o estoque inicial em vez do estoque médio). Isso tem levado a empresa a ter certa cautela nas quantidades do produto incluídas nos pedidos realizados. A dificuldade do proprietário está em equacionar os pedidos de barracas, de forma a atingir um nível de satisfação que considera aceitável – pelo menos 95% – e tentar maximizar o seu lucro. As demandas diárias do produto, em termos de probabilidades de ocorrência, baseadas em dados históricos dos últimos meses, são:

Demanda de barracas	11	12	13	14	15	16	17	18	19
Probabilidade	7%	9%	14%	20%	17%	13%	9%	7%	4%

Para cada pedido, é cobrado um frete de R\$ 150. O prazo de entrega dos pedidos também é variável e apresenta a seguinte distribuição de probabilidades:

Prazo de entrega dos pedidos (em dias)	2	3	4
Probabilidade	25%	55%	20%

No momento, o estoque é de 30 unidades e não existem pedidos a serem recebidos (pendentes). Você foi contratado pela empresa para elaborar um modelo de

administração de estoques (ou seja, definir o ponto e a quantidade de ressuprimento), objetivando maximizar o lucro e respeitando o nível mínimo de serviço (atendimento da demanda) exigido.

- a. Simule a experiência para 100 dias, alterando as variáveis referentes ao modelo de administração de estoques (ponto e quantidade de ressuprimento) e buscando atingir os objetivos da empresa. Como se comportam os componentes do lucro e o atendimento da demanda em face das mudanças nas variáveis citadas?
 - b. O que acontece com a busca da solução ótima, se o valor do frete passa a ser R\$ 250,00?
 - c. Refaça a simulação mantendo o valor do frete original, mas mudando o valor do aluguel diário do espaço para armazenar as barracas para R\$ 3,00 por unidade.
 - d. Elabore um relatório final com as suas recomendações.
3. Uma empresa comercial necessita dimensionar a capacidade de sua frota de entrega visando satisfazer certo nível de serviço. As entregas da empresa são realizadas apenas em dias úteis (adota-se a premissa de que o prazo normal – que ocorre quando existe capacidade de entrega – é de um dia útil) e as necessidades são aleatórias, com distribuição de probabilidades suposta uniforme, porém discreta, dependendo do dia da semana. A tabela a seguir descreve as necessidades de entrega mínima e máxima, de acordo com o dia da semana.

	Segunda	Terça	Quarta	Quinta	Sexta
Entrega mínima	21	10	11	11	11
Entrega máxima	45	22	23	23	23

O objetivo é estudar a distribuição do prazo de entrega dos produtos em função da capacidade diária de entrega, que poderá ser de 20, 22, 24, 26, 28 ou 30 volumes. Para isso, deseja-se estimar a média e o desvio-padrão desse prazo para cada uma das capacidades de entrega consideradas. Depois faça uma comparação entre as possíveis capacidades de entrega.

REFERÊNCIAS

- BOWERSOX, D.; CLOSS, D. *Logística empresarial: o processo de integração da cadeia de suprimento*. São Paulo: Atlas. 2001.
- CHWIF, L. Simulação aplicada à Logística. *Revista Mundo Logística*. Rio de Janeiro, nº 6 set./out.
- GOMES, C.; RIBEIRO, P. *Gestão da cadeia de suprimentos integrada à Tecnologia da Informação*. São Paulo: Thomson. 2004.
- JOHNSON, J.; WOOD, D. *Contemporary logistics*. New Jersey: Prentice Hall. 1996.
- MOORE, J. H.; WEATHERFORD, L. R. *Tomada de decisão em administração com planilhas eletrônicas*. 6. ed. Porto Alegre: Bookman. 2005.
- SHAMBLIN, J. E.; STEVENS JR. G. T. *Pesquisa operacional: uma abordagem básica*. São Paulo: Atlas, 1979.

Abordando probabilisticamente um problema com decisões sim ou não: um estudo de determinação de mix de projetos de consultoria

Carlos Alberto de Farias

MADE/UNESA

Marco Aurélio Carino Bouzada

MADE/UNESA

Objetivos de aprendizagem do capítulo

- Consolidar o aprendizado sobre modelagem e resolução de problemas segundo as abordagens vistas nos três primeiros capítulos do livro.
- Aprender a abordar problemas em condições de incerteza de forma não determinística, mas sem considerar (e precisar de) as probabilidades de ocorrência dos eventos.
- Saber atacar um problema de decisão por meio de diferentes abordagens: Programação Linear, Critérios não probabilísticos, Árvore de Decisão e Simulação de Monte Carlo.
- Verificar que tipo de conclusões cada uma dessas abordagens é capaz de fornecer.
- Entender as diferenças, vantagens, desvantagens e aplicabilidade de cada uma dessas abordagens.
- Visualizar um quadro comparativo entre essas abordagens de modo a saber identificar qual a abordagem mais adequada em vista das características do problema a ser atacado.

INTRODUÇÃO

O ambiente de tomada de decisão é dependente de quem decidirá e das informações que se tem a respeito do problema. Podemos identificar três ambientes no processo decisório: o de certeza, em que o tomador de decisão (administradores de empresas ou gerentes) identifica com certeza os efeitos de todas as alternativas e decisões; o de incerteza, em que o tomador de decisão não é capaz de atribuir as probabilidades de ocorrências aos vários resultados das alternativas identificadas; e, por fim, o ambiente

de risco, quando o tomador de decisão é capaz de atribuir probabilidades de ocorrência aos cenários.¹

Apesar da enorme contribuição das técnicas quantitativas aplicadas na solução de problemas gerenciais, em algumas situações elas têm potencial limitado. Uma decisão precisa ser tomada sempre que estamos diante de um problema que possui mais de uma alternativa para sua solução. Num modelo de processo decisório, é necessário analisar e identificar a situação e o problema, desenvolver e comparar as alternativas, classificar os riscos de cada uma delas, escolher a melhor e executar (implantar) e avaliar (validar ou não o processo utilizado), e para isso necessitamos de um modelo.²

Algumas situações se apresentam para a tomada de decisão, como: decisão em condições de certeza, que ocorre quando a decisão é assumida com pleno conhecimento de todos os estados da natureza; decisão em condições de risco, que ocorre quando são conhecidas as probabilidades de cada estado, podendo variar entre 0% e 100% para cada estado; decisão em condições de incerteza, que ocorre quando não se obteve total ou parcialmente os estados conhecidos e a probabilidade associada a elas; decisão em condições de competição ou conflito, quando há dois ou mais decisores envolvidos e o resultado depende da escolha de cada um dos decisores. Nestes três últimos casos, a decisão é considerada como não determinística.³

O aspecto não determinístico pode resultar de uma fonte de imprecisão causada pela dificuldade de avaliar as ações sob a influência de um ou mais critérios, pela indeterminação dos métodos de avaliação dos resultados, pela classificação ou opiniões ambíguas ou ainda por eventos probabilísticos. O problema é que, quando existe a incerteza, frequente no mundo empresarial, algumas técnicas não são adequadas. A solução determinística (sem considerar as incertezas) para um problema pode ser bastante desfavorável para a situação gerencial com a qual se depara a empresa.

Existem diversas ferramentas quantitativas já consagradas para adotar uma abordagem determinística para resolução de problemas, dentre elas a Programação Linear (ou o seu caso particular, a Programação Binária).⁴ Porém, em um contexto real das organizações, as incertezas existem, como, por exemplo, quanto à rentabilidade de um fundo de ações, e neste caso não deveria ser utilizado este tipo de abordagem.

Tomemos a situação de uma empresa que tem atualmente cinco propostas para desenvolvimento de projetos, cada uma com um percentual de expectativa de se tornar um contrato, baseado na experiência dos profissionais do setor comercial da empresa.

A empresa trabalha hoje com 33 analistas e 87 programadores, sendo que 21 analistas e 61 programadores estão alocados em projetos atuais e 12 analistas e 26 programadores estão terminando seus projetos. Cada proposta apresentada pela empresa para se tornar um contrato tem as características mostradas na [Tabela 4.1](#).

Quais propostas a empresa deveria desenvolver para obter maior lucro total, visto que a disponibilidade de recursos (analistas e programadores) é razoavelmente escassa?

1. Gomes; Gomes; Almeida (2009).

2. Gomes; Gomes; Almeida (2009).

3. Gomes; Gomes; Almeida (2009).

4. Hillier (2006); Medeiros (2008).

TABELA 4.1 Características de cada proposta

Proposta	Perspectiva de se tornar contrato	Recursos necessários		Lucro (em R\$)	Qtd Horas
		Analistas	Programadores		
P1	50%	5	8	780.000	17.350
P2	100%	4	6	450.000	10.080
P3	10%	5	10	900.500	20.010
P4	30%	6	12	1.633.000	36.288
P5	60%	3	5	303.500	6.720

Se a empresa tivesse certeza de que todas as propostas lançadas seriam aceitas, transformando-se em projetos, a Programação Binária⁵ poderia ser utilizada e a solução se tornaria trivial. Neste tipo de situação – ausência de incerteza – o uso de aplicações como a Programação Binária – que utiliza variáveis de decisão do tipo 0 ou 1 – pode desempenhar um papel particularmente importante na busca da solução. Inicialmente, então, o problema será tratado como se fosse determinístico, através da Programação Binária.

Para obtermos a solução ótima para o modelo de escolha da melhor proposta (ainda sem considerarmos a incerteza quanto à efetivação dos contratos), utilizamos a opção de restrição binária do Solver do Excel. No caso específico do problema, definimos os três elementos básicos para o modelo:

- Variáveis de decisão: fazer ou não cada proposta.
- Variável-objetivo: lucro total obtido, que deve ser maximizado.
- Restrições: recursos humanos exigidos/disponíveis (analistas e programadores).

Utilizando a Programação Binária, ainda partindo da premissa de que todas as propostas vão se transformar em contratos, a solução do problema é obtida a partir da aplicação do Solver. Neste caso, não há nenhum risco envolvendo a transformação das propostas em contrato. Pode-se observar que, quando não há nenhuma incerteza, podemos escolher os projetos 3 e 4 que, além de tudo, são os de maior lucro esperado.

A inadequação dessa abordagem nada tem a ver com o cálculo para descobrirmos, dentre as cinco propostas, quais deverão ser trabalhadas. O que falta nesse estudo são as incertezas. Em um contexto organizacional real, não é possível saber se as propostas vão se transformar em contrato: apenas se pressupõe a probabilidade de cada um deles se concretizar.

A solução para esse problema determinístico (sem a presença de riscos) pode ser bastante desfavorável para a situação gerencial com a qual se depara a empresa. Explica-se: na verdade, é pouco provável (10%) que a proposta 3 se transforme em projeto. Se a empresa elaborar apenas as propostas 3 e 4 (que é o que sugere a solução obtida

5. Moore; Weatherford (2005); Hillier (2006); Leão *et al.* (2006).

através da abordagem primária), é bem possível que o projeto 3 não se concretize, ficando a empresa com recursos ociosos, que poderiam ter sido utilizados, por exemplo, no projeto 1, apenas um pouco menos rentável, mas com probabilidade de efetivação muito maior. Por outro lado, não podem ser elaboradas propostas demais, porque então passa a existir o risco de não haver pessoal suficiente para atender a todos os projetos.

Evoluindo, então, para uma abordagem mais realista – já que a metodologia simples e trivial da Programação Binária (que utiliza um modelo determinístico) não se aplica adequadamente nesse caso em que temos risco e incertezas – precisaremos utilizar um modelo que os considere, necessitando ser tratado através de uma técnica não determinística como: os critérios Maximax, Maximin e Minimax;⁶ a Árvore de Decisão;⁷ ou a Simulação.⁸

Dessa forma, a proposta principal deste capítulo consiste em analisar os resultados obtidos pela aplicação de diferentes abordagens que considerem incertezas e riscos ao problema de escolha das propostas, de forma a mostrar que tipo de conclusões essas abordagens são capazes de obter em casos normalmente tratados de forma determinística, a fim de se constituir um quadro comparativo.

A seguir é apresentada uma breve revisão dos conceitos acerca de Programação Binária, Critérios não probabilísticos, Árvore de Decisão e Simulação.

PROGRAMAÇÃO BINÁRIA

Uma área de aplicação muito importante da Programação Linear é aquela que envolve problemas com uma série de decisões sim ou não, fazer ou não fazer. Tais variáveis são chamadas variáveis binárias e os problemas de Programação Inteira contendo apenas variáveis binárias são denominados Programação Binária (ou PI 0-1).⁹

As variáveis binárias, 0 ou 1, desempenham um papel particularmente importante, já que possibilitam incorporar instruções condicionais a decisões do tipo sim ou não, de forma semelhante ao Excel. Estas decisões são, por vezes, chamadas de dicotômicas, em um modelo de otimização.¹⁰

Os gerentes frequentemente precisam enfrentar decisões do tipo sim ou não. Portanto, a Programação Inteira Binária (PIB) é amplamente usada para ajudar nessas decisões. Elas possibilitam lidar com um problema com cuja formulação natural é difícil de se trabalhar e reformulá-lo como um problema de PI pura ou mista.¹¹

CRITÉRIOS NÃO PROBABILÍSTICOS

O critério Maximax se baseia em uma visão otimista do problema. Escolhido determinado modelo, supõe-se que ocorrerá o melhor evento possível. Neste critério, o problema decisório pode levar-nos à solução ótima, em que uma alternativa será escolhida pela

6. Gomes; Gomes; Almeida (2009); Corrar; Theóphilo (2004); Moore; Weatherford (2005).

7. Corrar; Theóphilo (2004); Moore; Weatherford (2005).

8. Hertz (1980); Corrar; Theóphilo (2004); Moore; Weatherford (2005); Mendes *et al.* (2006).

9. Hillier (2006).

10. Moore; Weatherford (2005).

11. Hillier (2006).

maximização da função-objetivo. Este método é indicado quando buscamos a alternativa com o melhor desempenho de maximização. O critério Maximax carrega consigo uma visão de mundo muito otimista, ou seja, deve-se escolher o melhor resultado de cada alternativa, e dentre eles o melhor dos melhores resultados.¹² O método pode ser definido como critério de decisão otimista, que tem por objetivo encontrar o melhor resultado possível, a alternativa de decisão que maximizará o máximo resultado.¹³

O critério Maximax avalia cada decisão pelo que melhor pode acontecer se o gerente tomar essa decisão. Neste caso eles avaliam cada decisão pelo máximo retorno possível associado àquela decisão, porém não se deve confundir a decisão com o estado de natureza que produz o resultado ótimo.¹⁴

Já o critério Maximin se baseia em uma visão pessimista do problema. Supõe-se que, escolhido determinado modelo, ocorrerá o pior evento possível. A alternativa será escolhida como aquela que tem a melhor entre as piores opções de todas as alternativas. Deve-se determinar o lucro mínimo para cada alternativa e, em seguida, escolher a alternativa com o maior lucro mínimo.

O critério Maximin, tido como extremamente conservador ou pessimista ao tomar decisões, tem por objetivo encontrar a decisão que maximizará o mínimo resultado entre vários possíveis, ou seja, ele indicará a alternativa que produz o maior resultado entre os mínimos encontrados.¹⁵

Ao usar esse critério, o tomador de decisões deve ficar alerta para o que significa máximo ou mínimo, dependendo de como são expressos os resultados da matriz de decisão.¹⁶

Finalmente, a estrutura do critério Minimax é baseada inicialmente desenvolvendo-se uma tabela de perdas de oportunidade, a seguir encontrando-se a oportunidade máxima de perdas em cada alternativa e selecionando-se entre elas a alternativa que apresenta a menor perda.¹⁷

Moore e Weatherford (2005, p. 385) destacam que “arrependimento é sinônimo de ‘custo de oportunidade’ de não tomar a melhor decisão para um dado estado de natureza”. Em uma tomada de decisão, o gerente gostaria de tomar uma decisão que minimizasse o arrependimento, mas não sabe que estado de natureza acontecerá. Se o gerente não conhece a probabilidade, o que se sugere é usar o critério Minimax conservador, selecionando a decisão que faz o melhor do pior caso, ou seja, a decisão que tem o menor arrependimento máximo.

ÁRVORE DE DECISÃO

A Teoria da Decisão é um conjunto de técnicas quantitativas que tem por objetivo ajudar o gerente a sistematizar o problema e solucioná-lo. Não há solução de um problema sem critério, e é neste sentido que a Teoria da Decisão baseia-se em

12. Moreira (2007).

13. Corrar; Théophilo (2004).

14. Moore; Weatherford (2005).

15. Corrar; Théophilo (2004); Moore; Weatherford (2005).

16. Gomes; Gomes; Almeida (2009); Moreira (2007).

17. Corrar; Théophilo (2004).

critérios preestabelecidos, havendo sempre espaço para novos critérios e novas contribuições.¹⁸

Para construir uma Árvore de Decisão é necessário desenhar um gráfico de modo que as informações adicionais obtidas pelo tomador de decisão, no transcorrer do processo, possam ser incluídas como novas variáveis. A construção da Árvore de Decisão respeita a sequência utilizada na leitura da Tabela de Decisão, em que parte das alternativas leva em conta os cenários e culmina com o indicador monetário que definirá a escolha do tomador de decisão, que é o resultado que retornará em decorrência da decisão.¹⁹

Existem dois critérios usualmente utilizados no ambiente de risco: o Valor Monetário Esperado (VME), que maximiza os resultados esperados e que, para cada alternativa de decisão, sua apuração é possível pelo cálculo da média dos valores de seus diversos resultados, ponderada pelas probabilidades de ocorrência dos eventos; e a Perda de Oportunidades Esperada (POE), que minimiza as perdas e refere-se à diferença entre o retorno ótimo e o retorno recebido.²⁰

Para determinar o valor esperado para cada alternativa e escolher a alternativa que tiver o melhor valor esperado, deve-se atribuir uma probabilidade a cada acontecimento (mutuamente exclusivos), calcular os valores esperados de cada ação, multiplicando cada valor consequente pela correspondente probabilidade, somar esses produtos e escolher a ação cujo valor esperado seja o máximo.²¹

Na análise de uma decisão cujas possíveis consequências são valores monetários relativamente próximos, o VME pode ser utilizado como critério de decisão quando da escolha de uma dentre várias alternativas. No entanto, nos casos em que os valores monetários das várias alternativas são substancialmente diferentes, como numa decisão estratégica, o VME não mais reflete necessariamente como um tomador de decisão avalia uma estratégia de ação, pois não leva em conta a sua possível aversão a grandes perdas.²²

Uma alternativa ao VME é a Utilidade, que consiste na atratividade dos possíveis resultados de decisão, já que a maioria das pessoas é avessa ao risco. A análise decisória lida com esse comportamento de aversão ao risco. A função Utilidade é definida como uma medida de satisfação, diferente para cada tomador de decisão, e que substitui o VME no momento de considerar o valor equivalente de um nó de decisão.²³

Uma Árvore de Decisão deve ser complexa o suficiente para incorporar todos os elementos-chave e valores que são importantes para os pacientes e, ao mesmo tempo, simples o suficiente para ser compreensível e operacional.²⁴

SIMULAÇÃO

A origem da Simulação é muito antiga. Os chineses já a utilizavam em seus jogos de guerra há 5 mil anos e os povos prussios utilizaram esses jogos no final do século XVIII para auxiliá-los no treinamento militar de suas tropas. Durante a

18. Moreira (2007).

19. Corrar; Theóphilo (2004).

20. Corrar; Theóphilo (2004).

21. Ferreira (2004).

22. Vatter *et al.* (1978).

23. Moore; Weatherford (2005).

24. Campolina; Ciconelli (2006).

Segunda Guerra Mundial, o matemático John Von Neumann criou o conceito denominado Simulação de Monte Carlo, que consistia na simulação direta de problemas probabilísticos. Atualmente com o desenvolvimento dos recursos computacionais, esse método é usado em diversas áreas, desde a simulação de fenômenos físicos complexos até menos complexos, como a simulação do resultado de loterias.²⁵

Na Simulação, procura-se obter informações sobre os relacionamentos entre as variáveis que compõem o sistema no sentido de prever seus futuros desempenhos sob novas condições. O objetivo é criar um ambiente no qual a informação sobre ações alternativas possíveis é conseguida através da experimentação.²⁶

A DETERMINAÇÃO DO MIX IDEAL DE PROJETOS ATRAVÉS DE ABORDAGENS NÃO DETERMINÍSTICAS

Além das variáveis utilizadas na abordagem do modelo binário (decisão de fazer ou não cada proposta, lucro esperado e restrições de pessoal), incluímos, nas análises probabilísticas (ver *Árvore da Decisão e Simulação*, a seguir), a variável expectativa de transformar as propostas em contratos, estimada com a ajuda da experiência dos gerentes, já apresentadas na [Tabela 4.1](#).

Critérios Não Probabilísticos

O problema de escolha das propostas foi resolvido segundo os três critérios: Maximax, Maximin e Minimax. Para resolver esse problema, foram elencadas as 31 soluções possíveis criadas pela combinação das cinco propostas, e os possíveis resultados para cada uma delas em função das propostas se transformarem ou não em contrato.

De acordo com o critério Maximax, foi escolhido o melhor resultado possível; com o critério Maximin, o resultado que produz o maior resultado entre os mínimos encontrados; e com o critério Minimax, a alternativa que minimizará a maior perda de oportunidade.

Vale ressaltar que, segundo esses três critérios não probabilísticos, conforme o nome sugere, as probabilidades não estão contempladas; levou-se em conta apenas que o projeto pode ou não ser feito. Além disso, foi considerada uma multa para as situações em que a empresa não dispuser de recursos humanos suficientes para lidar com todas as propostas aprovadas.

Critério Maximax

As variáveis de decisão dependem exclusivamente do tomador de decisão. Ele pode decidir fazer somente a proposta 1, ou só a 2 ou ainda optar por qualquer uma das 31 possibilidades de solução. Supondo que tenha escolhido fazer a 3 e a 4, o que vai acontecer com o lucro se somente a 1 viesse a ser aprovada? Nesse caso, a proposta

25. Corrar; Theóphilo (2004).

26. Moore; Weatherford (2005).

1 não foi feita e não pode ser aprovada, e as outras duas foram feitas, mas não aprovadas, gerando um lucro nulo. Se a proposta 3 for aprovada, a empresa vai ganhar o lucro referente a ela. A melhor possibilidade é fazer as cinco propostas (1, 2, 3, 4 e 5, ou seja, C12345 x 12345), o que demanda 23 analistas e 41 programadores, e portanto não há recursos suficientes para fazer todas as propostas, pois o total de recursos disponíveis é de 23 programadores e 12 analistas.

Quanto seria perdido pelo desgaste comercial de assumir fazer as cinco propostas e não ter condição de executá-las? A empresa vive um momento instável, com muitas ocorrências, e por este motivo, para resolver este problema, foi criada uma variável que definimos como “% residual de lucro no caso de falta de recursos”, que arbitramos como 50% (ou seja, é como se, em cima do lucro, fosse aplicada uma “Multa” de $100\% - 50\% = 50\%$) para o caso de a proposta tornar-se contrato e não haver recursos suficientes para executar o projeto. Foi criada uma fórmula aplicada em todas as células para definir de quanto será o lucro. Assim, tomemos como exemplo a relação “Fazer as propostas” 1 e 3 x “Propostas que vão virar contrato” 2345. Se a proposta 1 foi feita e não foi aprovada, na célula 2345 é somado apenas o lucro da proposta 3. Aplicando o valor de Multa de 50%, o lucro máximo dos máximos foi de R\$ 2.533.500,00, para as alternativas C34, C134, C234, C345, C1234, C1345, C2345 ou C12345. Para efeitos comparativos, processamos, respectivamente, os resultados para o caso de a Multa ser 70% (cenário pessimista) e 30% (cenário otimista).

Os resultados encontrados com o valor da Multa de 70% (cenário pessimista) são os mesmos encontrados para valor da Multa de 50%, porém, os resultados apresentados para o valor da Multa de 30% (cenário otimista) tiveram uma mudança significativa, mostrando que apenas a alternativa C12345 terá o valor máximo dos máximos, ou seja, R\$ 2.846.900,00.

Critério Maximin

Os parâmetros são os mesmos considerados no critério Maximax. Novamente utilizamos a Multa de 50%, e neste caso o valor máximo dos mínimos apontou para fazer, indistintamente, qualquer conjunto de propostas, todos com lucro de R\$ 0,00. Em razão da natureza pessimista desta abordagem, não faz sentido elaborar a análise de sensibilidade no que se refere à variável Multa, já que, se tudo der errado, nenhuma proposta virará contrato e o valor da Multa não estará influenciando a decisão.

Critério Minimax

Os parâmetros e os resultados intermediários são os mesmos encontrados para o critério Maximax. Novamente utilizamos a Multa de 50%, e neste caso o valor máximo dos mínimos apontou para fazer as propostas 1, 3 e 4 com lucro de R\$ 876.750,00.

Para efeito de comparação, também foram feitas análises para a Multa com 70% e 30%, respectivamente, e neles os resultados diferem do resultado apresentado para a Multa normal de 50%, sendo, respectivamente, R\$ 1.533.500,00 para fazer as propostas 3 e 4, e R\$ 445.050,00 para as propostas P1, P2, P3 e P4.

Para facilitar o entendimento, o [Quadro 4.1](#) mostra o resumo das melhores situações encontradas por cada critério/cenário.

QUADRO 4.1 Quadro comparativo Maximax x Maximin x Minimax

	Maximax		Maximin		Minimax	
	Fazer as propostas	Máximo	Fazer as propostas	Máximo	Fazer as propostas	Mínimo
Realista						
	C 34	R\$ 2.533.500	Todas	R\$ -	C 134	R\$ 876.750
	C 134	R\$ 2.533.500				
	C 234	R\$ 2.533.500				
	C 345	R\$ 2.533.500				
	C 1234	R\$ 2.533.500				
	C 1345	R\$ 2.533.500				
	C 2345	R\$ 2.533.500				
	C 12345	R\$ 2.533.500				
	Máximo dos Máximos		Máximo dos Mínimos	R\$ -	Mínimo dos Máximos	R\$ 876.750
Otimista						
	C 12345	R\$ 2.846.900	Todas	R\$ -	C 1234	R\$ 445.050
	Máximo dos Máximos	R\$ 2.846.900	Máximo dos Mínimos	R\$ -	Mínimo dos Máximos	R\$ 445.050
Pessimista						
	C 34	R\$ 2.533.500	Todas	R\$ -	C 34	R\$ 1.533.500
	C 134	R\$ 2.533.500				
	C 234	R\$ 2.533.500				
	C 345	R\$ 2.533.500				
	C 1234	R\$ 2.533.500				
	C 1345	R\$ 2.533.500				
	C 2345	R\$ 2.533.500				
	C 12345	R\$ 2.533.500				
	Máximo dos Máximos		Máximo dos Mínimos	R\$ -	Mínimo dos Máximos	R\$ 1.533.500

Árvore de Decisão

Foi feita uma Árvore de Decisão contemplando as 31 possíveis soluções (as mesmas da seção anterior), porém, nesse caso, foram utilizadas as probabilidades (expectativas) para considerar a chance de uma proposta virar ou não contrato. Em cada uma das soluções foram consideradas as probabilidades de as propostas se transformarem em contrato para avaliar o resultado financeiro. A mesma multa mencionada na seção anterior foi considerada para esta metodologia. A Árvore de Decisão foi desenvolvida inicialmente segundo o critério do Valor Esperado. Fazer as propostas é uma decisão do tomador de decisão. A árvore foi desenvolvida de modo a comportar a decisão de fazer uma proposta, duas propostas, três propostas, quatro propostas ou ainda de fazer as cinco propostas. Para cada possível decisão, verificamos quais as probabilidades de as propostas virarem contrato. Uma das possibilidades é nenhuma virar contrato. Isto não significa que não foi feita nenhuma proposta, e sim que, nesse caso, nenhuma vai virar contrato.

Se todas as propostas fossem feitas e todas virassem contrato, a empresa ganharia R\$ 4.067.000,00, porém, não haveria recursos humanos suficientes para o desenvolvimento de todos os projetos, causando atraso e desgaste comercial, que foi considerado a partir da aplicação de uma “Multa” que reduz (em um determinado percentual) o lucro obtido. Este percentual foi fixado inicialmente em 50% e a Multa foi aplicada somente nos casos específicos em que não houve recursos.

Assim, a Árvore de Decisão, utilizando o software TreePlan, escolhe a solução com melhor Valor Esperado. O nó de decisão (fazer as cinco propostas) foi escolhido como o melhor resultado, com Valor Esperado de R\$ 1.283.349,00.

Segundo a execução do software TreePlan, dado que haja 50% de chance de a empresa ganhar R\$ 780.000,00, e 50% de chance de ela ganhar R\$ 0,00, o Valor Esperado calculado é de R\$ 390.000,00, mostrado na [Figura 4.1](#). Para o cálculo da árvore através da Função de Utilidade, foi coletado com o tomador de decisão o seu Nível de Aversão ao Risco, tendo sido indagado qual seria a equivalente-certeza para esta situação de risco descrita anteriormente. O valor de R\$ 250.000,00 apareceu como resposta e foi aplicado na função Atingir Meta de forma a se obter o valor de Tolerância ao Risco de R\$ 494.827,00, conforme ilustra a [Figura 4.2](#).

A árvore também nos mostrou algumas informações importantes. Se forem feitas quatro propostas, o melhor seria fazer as propostas 1, 2, 4 e 5, com lucro de

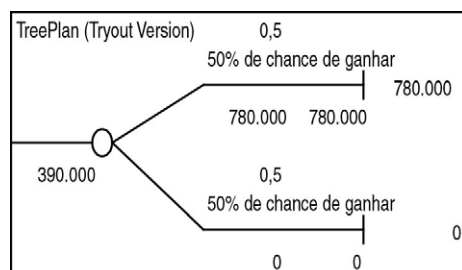


FIGURA 4.1 Nó de evento – Valor Esperado.

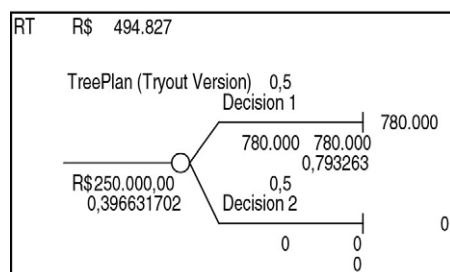


FIGURA 4.2 Nó de evento – Tolerância ao Risco.

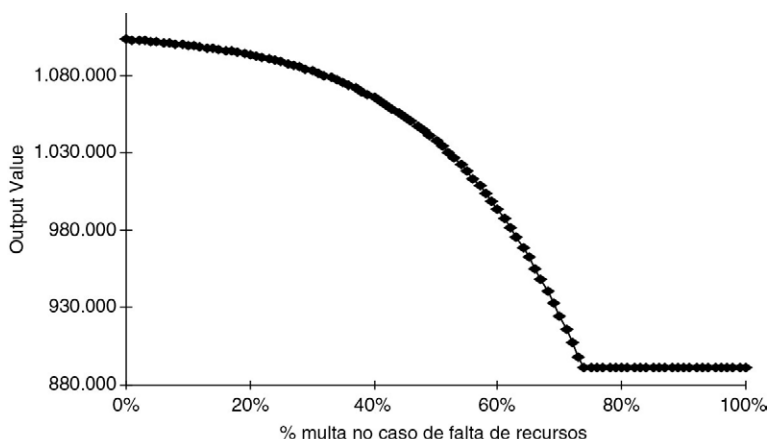


FIGURA 4.3 Análise de Sensibilidade – impacto da Multa sobre o valor da árvore.

R\$ 930.898,00. Se for escolhido fazer três propostas, a melhor opção seriam a 1, a 2 e a 5, com lucro de R\$ 891.905,00. Caso se decidisse por duas propostas, a melhor opção consistiria nas propostas 1 e 2, com lucro de R\$ 700.000,00. Por último, se a opção é fazer apenas uma proposta, a melhor seria a número 2, com lucro de R\$ 450.000,00.

Com a Árvore de Decisão construída, procedeu-se a análise de sensibilidade com as variáveis mais importantes. Primeiro foi trabalhada a Multa, variando de 0% a 100%, de 1% em 1%. O *output* é a solução ótima, em que concluímos que de 0% a 73% o valor da Multa não importa para efeitos de determinação da melhor solução e podem ser feitas as cinco propostas. Porém ao aplicarmos o percentual de Multa de 74% o melhor seria fazer três propostas, sendo o conjunto 1, 2 e 5 o melhor deles, com lucro de R\$ 891.509,00, valor mostrado na Figura 4.3, quando aplicamos a análise de sensibilidade para o valor final da árvore.

Percebe-se na Figura 4.3 que, à medida que a Multa aumenta, o valor da solução com cinco propostas vai diminuindo; mas o valor da Multa passa a ser irrelevante quando a solução ótima consiste de apenas três propostas, pois nunca faltarão recursos neste caso.

A Árvore de Decisão também foi analisada com cenários pessimista e otimista. Em um cenário mais pessimista, utilizamos as perspectivas de transformação das propostas em contrato como sendo: 40%; 80%; 5%; 20%; e 50%. A partir desta situação, para uma Multa de 50%, encontramos um lucro de R\$ 750.164,00, sendo a melhor solução fazer as cinco propostas.

A Análise de Sensibilidade do impacto da Multa na decisão ótima mostra que de 0% a 90% o valor da Multa não importa e podem ser feitas as cinco propostas, porém, ao aplicarmos o percentual de Multa de 91%, o melhor seria fazer três propostas, sendo o conjunto 1, 2 e 5 o melhor deles, com lucro de R\$ 639.198,00, quando aplicamos a Análise de Sensibilidade para o valor final da árvore.

A Análise de Sensibilidade do impacto da Multa foi tratada também com o cenário otimista (probabilidades de efetivação dos contratos iguais a 60%, 100%, 15%, 40% e 70%). O Quadro 4.2 mostra o resumo das melhores soluções encontradas pelo método da Árvore de Decisão (abordagens do Valor Esperado e da Função de Utilidade) em diferentes cenários.

Simulação

Nesta abordagem, também foi incluída a variável expectativa de transformar as propostas em contratos. Neste caso, o lucro total dependeu efetivamente de uma combinação

QUADRO 4.2 Quadro comparativo – Árvore de Decisão

	Multa Aplicada	Fazer 'n' Propostas	Melhores Propostas	Lucro
Valor Esperado	50%	5	1, 2, 3, 4, 5	R\$ 1.283.349,00
Função Utilidade	Análise de Sensibilidade			
Realista				
Cenário 1	0 – 73%	5	1, 2, 3, 4 e 5	R\$ 1.103.547,00 a R\$ 898.056,00
Cenário 2	74% ou mais	3	1, 2, 5	R\$ 891.509,00
Pessimista				
Cenário 1	0 – 90%	5	1, 2, 3, 4 e 5	R\$ 766.253,00 a R\$ 639.912,00
Cenário 2	91% ou mais	3	1, 2, 5	R\$ 639.198,00
Otimista				
Cenário 1	0 – 71%	5	1, 2, 3, 4 e 5	R\$ 1.268.477,00 a R\$ 964.963,00
Cenário 2	72% ou mais	3	1, 2, 5	R\$ 961.258,00

específica de um grande número de variáveis diferentes. A simulação foi necessária para verificar se cada proposta elaborada iria se transformar efetivamente em projeto. Tal verificação ocorreu a partir da geração de números aleatórios que, em conjunto com as probabilidades de sucesso na efetivação de cada proposta, foi capaz de simular se cada contrato se efetivou, para cada experimento realizado.

O trabalho procurou simular o que aconteceria com o lucro total e com adequação do pessoal exigido ao pessoal disponível em cada uma de todas as possíveis soluções para o problema de decidir quais propostas deverão ser feitas. São 31 soluções possíveis: fazer todas as propostas (P1, P2, P3, P4, P5), fazer somente uma (P1; P2; P3; P4 ou P5), fazer duas, três ou quatro propostas. Foram realizados 5 mil experimentos para cada uma das soluções.

Para cada proposta foi sorteado um número aleatório, em cada simulação. Em função deste número aleatório e da probabilidade de efetivação de cada proposta, foi verificado em cada uma das 5 mil simulações se a proposta iria virar projeto ou não, o que está representado pelos valores 1 ou 0 na planilha, respectivamente.

Tratando apenas os projetos que vão ser efetivamente feitos, foram calculadas as quantidades de analistas e programadores necessários para a sua execução em função das necessidades básicas de cada projeto. Estes valores foram comparados aos recursos disponíveis, para verificar se cada conjunto de recursos necessários é viável. Se ambos os conjuntos foram viáveis, a viabilidade geral (célula “OK geral”) recebeu o valor 1. A célula “% de OK” contabilizou em quantas das 5 mil simulações os recursos disponíveis puderam atender a demanda por eles. Assim, uma medida de risco pode ser construída ($\text{risco} = 100\% - \% \text{ OK}$) para contemplar qual a probabilidade de tal conjunto de propostas feitas gerar um desgaste comercial por conta de não haver recursos humanos necessários para atender a demanda de todos os contratos efetivados.

O outro indicador de qualidade da solução foi o lucro total, função exclusivamente dos projetos que foram efetivados e do lucro esperado de cada um. Ao final da corrida de 5 mil simulações, foi possível ter uma ideia do lucro médio (ou esperado) da solução analisada.

Aqui é muito importante deixar claro uma premissa que foi utilizada na modelagem: mesmo que os recursos disponíveis não sejam suficientes para atender a demanda de todos os projetos, ainda assim o lucro esperado de cada um deles será considerado no cálculo do lucro total. Tal consideração é razoável, porque a expectativa é que todos os projetos aprovados venham a ser feitos em algum momento, a partir da liberação gradual dos recursos oriundos dos projetos que forem chegando ao fim. A realização não imediata do projeto, no entanto, pode gerar um desgaste comercial junto ao cliente, e portanto esse risco precisou ser monitorado através da medida de risco introduzida anteriormente.

A planilha apresentada na [Tabela 4.2](#) ilustra a simulação de parte dos 5 mil experimentos realizados para a solução C 2345. A partir da simulação de todas as soluções exequíveis, foi possível fazer um quadro sumarizando o lucro e o nível de risco previstos para cada uma.

Vale ressaltar que esses resultados são, na realidade, dinâmicos. Cada vez que a planilha é atualizada, os valores aleatórios são recalculados e novos resultados são obtidos. As ilustrações apresentadas consistem apenas em “fotografias” de uma dessas atualizações. Felizmente, com 5 mil simulações, os resultados são robustos, ou seja, não apresentam grandes variações entre as atualizações.

A Figura 4.4 mostra os mesmos resultados, porém, abordando a relação Lucro x Risco e indicando um ponto de corte para um nível máximo de risco aceitável (tendo a área comercial da empresa sinalizado ser interessante não haver mais do que 5% de probabilidade de haver desgaste comercial).

A linha pontilhada na Figura 4.4 indica o que seria um patamar desejável de risco máximo tolerável (2,5%). A partir da análise do quadro e das figuras, verificamos que o conjunto de propostas P1 + P2 + P5 não traz risco algum e fornece uma perspectiva de lucro de R\$ 1.000.000, aproximadamente, sendo esta, a princípio, uma boa opção.

Se apenas a análise minimizadora de risco for utilizada, os executivos poderão ir adiante, sem a preocupação de estarem correndo risco. Neste caso, o conjunto P1 + P2 + P5 ou o P2 + P3 + P5 poderiam ser escolhidos. Verificamos que as propostas P2 e P5 aparecem, nas duas composições, com 0% de risco.

Para executar todas as propostas, teríamos um nível de risco muito grande, de aproximadamente 29%. Comportando um pequeno risco (aproximadamente 2,5%), duas soluções aparecem: P1 + P3 + P4 e P3 + P4 + P5 (as propostas P3 e P4 aparecem

TABELA 4.2 Resultado (parcial) da Simulação para o caso C 2345

	P1	P2	P3	P4	P5	Totais
Fazer ou não	0	1	1	1	1	para
Expectativa	50%	100%	10%	30%	60%	100%
Lucro R\$	780.000	450.000	900.500	1.633.000	303.500	4.067.000
Analistas	5	4	5	6	3	23
Programadores	8	6	10	12	5	41
Total Recursos	13	10	15	18	8	
Recursos para						
Desenvolvimento dos Projetos						
Analistas				Programadores		
12				26		
Totais						
18				33		
% OK 80,3%						
Lucro Total Médio						
R\$ 1.212.993,90						

(Continua)

TABELA 4.2 Resultado (parcial) da Simulação para o caso C 2345 (*cont.*)

Simulação	ALEATÓRIOS					PROJETOS					ANALISTAS		PROGRAM.			Lucro
	1	2	3	4	5	P1	P2	P3	P4	P5	Utilizados	Ok?	Utilizados	Ok?	OK geral	Total
1	0,74	0,75	0,75	0,56	0,11	0	1	0	0	1	7	1	11	1	1	R\$ 753.500,00
2	0,73	0,96	0,16	0,09	0,25	0	1	0	1	1	13	0	23	1	0	R\$ 2.386.500,00
3	0,11	0,15	0,01	0,70	0,67	0	1	1	0	0	9	1	16	1	1	R\$ 1.350.500,00
4	0,68	0,65	0,74	0,52	0,13	0	1	0	0	1	7	1	11	1	1	R\$ 753.500,00
5	0,44	0,48	0,73	0,10	0,77	0	1	0	1	0	10	1	18	1	1	R\$ 2.083.000,00
6	0,72	0,53	0,67	0,50	0,67	0	1	0	0	0	4	1	6	1	1	R\$ 450.000,00
7	0,29	0,51	0,14	0,45	0,50	0	1	0	0	1	7	1	11	1	1	R\$ 753.500,00
8	0,34	0,75	0,62	0,80	0,51	0	1	0	0	1	7	1	11	1	1	R\$ 753.500,00
9	0,44	0,51	0,70	0,18	0,48	0	1	0	1	1	13	0	23	1	0	R\$ 2.386.500,00
10	0,82	0,78	0,04	0,31	0,07	0	1	1	0	1	12	1	21	1	1	R\$ 1.654.000,00

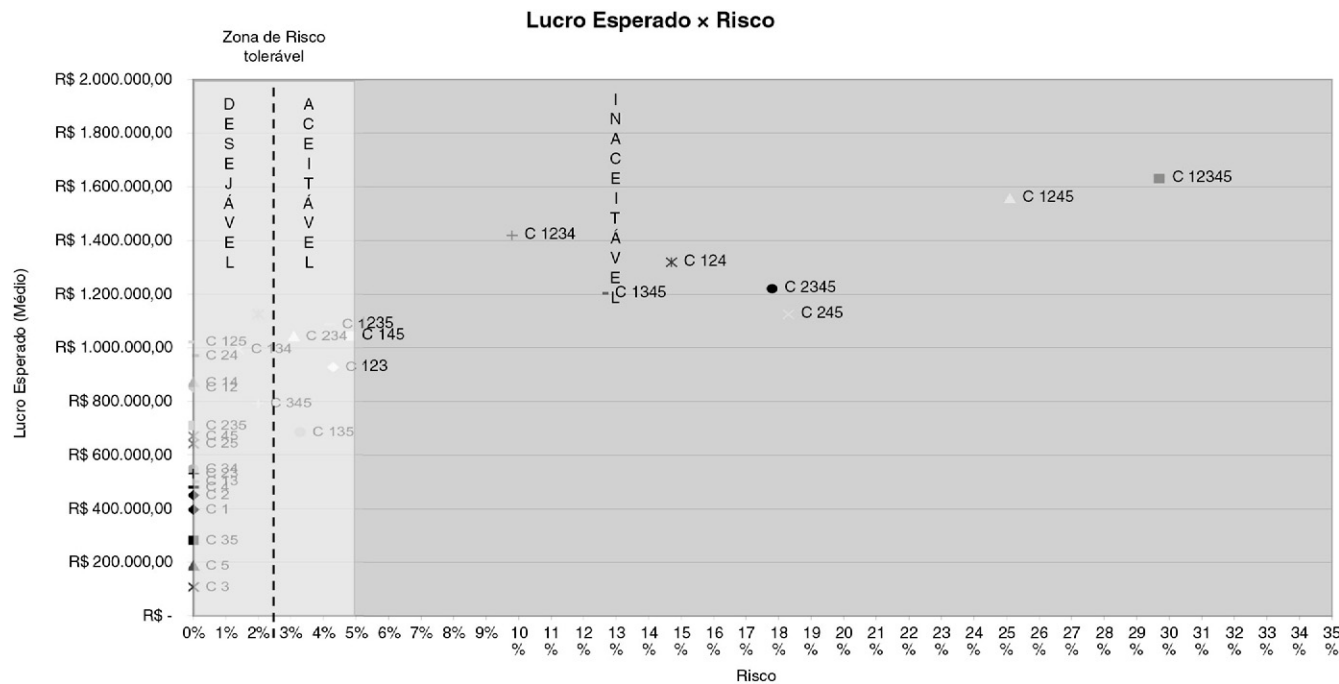


FIGURA 4.4 Lucro e risco de cada solução, com ponto de corte.

nas duas combinações). A primeira delas apresenta um lucro esperado de aproximadamente R\$ 1.000.000.

Ainda dentro do risco delimitado pela área comercial (5%), qualquer combinação que leve em conta a execução de três propostas pode ser recomendada, exceto os conjuntos $P1 + P2 + P4$ e $P2 + P4 + P5$. As soluções $P1 + P2 + P3$ e $P1 + P4 + P5$ estão fora da zona aceitável, porém, em algumas simulações, elas são consideradas.

Dentre as soluções aceitáveis com três propostas, as combinações $P1 + P2 + P5$ e $P2 + P3 + P5$ são as mais lucrativas (em torno de R\$ 1.000.000). Porém, a solução $P1 + P2 + P3 + P5$ também aparece na zona aceitável de risco (5%, aproximadamente), mas com um lucro esperado de cerca de R\$ 1.100.000.

Verificamos que, além de ser de interesse da empresa apresentar o maior número possível de propostas (por questões estratégicas), devemos fazer as propostas $P1$, $P2$, $P3$ e $P5$, dentro do nível de risco aceitável e obtendo a maior expectativa de lucro.

Como as probabilidades de as propostas se transformarem em contrato foram baseadas na experiência dos gerentes e são, portanto, subjetivas, convém realizar uma análise de sensibilidade a respeito do seu impacto.

Tratamos também essa abordagem em um cenário mais pessimista, em que as perspectivas de transformação das propostas em contrato poderiam ser, respectivamente: 40%; 80%; 5%; 20%; e 50%, e em um cenário mais otimista, no qual as perspectivas de transformação das propostas em contrato poderiam ser, respectivamente: 60%; 100%; 15%; 40%; e 70%.

O **Quadro 4.3** mostra uma comparação das melhores soluções encontradas pelo método de Simulação para os cenários realista, otimista e pessimista.

Análise Comparativa das Abordagens

Não Probabilísticas x Probabilísticas

Como vantagem das abordagens não probabilísticas (Maximax, Maximin, Minimax), podemos considerar que nelas não há dependência da subjetividade (e de uma eventual imprecisão) das probabilidades utilizadas na Árvore de Decisão e na Simulação. Se estas probabilidades estiverem erradas (elas são subjetivas), isto pode comprometer um pouco estas duas abordagens. Para as não probabilísticas, essa consideração não entra na composição.

De qualquer forma, trata-se de uma pequena vantagem; porém, em contrapartida, as abordagens não probabilísticas apresentam uma desvantagem porque estaremos tratando todas as propostas da mesma maneira, ou seja, todas as propostas podem ou não se transformar em contrato, quando, na realidade, algumas têm mais chance do que outras.

Comparação entre os três critérios não probabilísticos

Como desvantagem, a abordagem Maximax apresentou múltiplas soluções ótimas em todos os cenários. A abordagem Maximin simplesmente não funcionou para este problema.

A abordagem Minimax foi a melhor: apresentou apenas uma solução ótima para cada cenário e revelou certa sensibilidade em relação à Multa, ou seja, para cada valor da Multa, ela disponibilizou uma solução ótima diferente.

QUADRO 4.3 Quadro comparativo – Simulação

Realista	Solução	C 123	C 125	C 145	C 235	C 245	C 1235
	Lucro R\$	719.007,00	839.369,50	783.392,00	564.151,00	837.167,50	875.276,50
	Risco	1,40%	0,00%	2,10%	0,00%	7,80%	2,10%
Otimista	Solução	C 125	C 134		C 235	C 345	
	Lucro R\$	1.125.421,00	1.282.459,00		785.221,50	984.605,50	
	Risco	0,00%	4,40%		0,00%	3,80%	
Pessimista	Solução			C 145	C 234		C 1234
	Lucro R\$			783.392,00	780.942,44		1.011.647,50
	Risco			2,10%	1,40%		3,00%

Árvore de Decisão x Simulação

Podemos considerar como uma pequena desvantagem da Simulação o fato de a planilha ficar muito “pesada”. Além disso, ela não apresentou diretamente a solução ótima, diferentemente da Árvore de Decisão; ela apenas mostra o desempenho de cada uma das possíveis soluções. Mas como a quantidade de soluções possíveis não é muito grande (apenas 31), foi possível verificar o desempenho de todas elas na Simulação, o que não é uma desvantagem tão marcante.

A confecção da Árvore de Decisão é muito trabalhosa, e isso pode ser considerado uma pequena desvantagem; porém, como vantagem, ela não ficou muito “pesada”. Na Árvore de Decisão, foi possível contemplar o perfil de risco do tomador de decisão. Na Simulação isso não foi possível.

Podemos colocar como desvantagem da Árvore de Decisão o fato de que assumimos uma Multa para penalizar o lucro nos casos em que a quantidade de recursos disponíveis não foi suficiente para atender todas as demandas de projeto. A desvantagem deve-se ao fato de esta multa ter sido arbitrada.

Na Simulação, a análise do problema baseou-se na verificação da probabilidade da demanda por recursos superar a oferta. Por este motivo, o tratamento foi menos arbitrário quando comparado com o da Árvore de Decisão: nesta última, arbitramos a Multa; na Simulação, mensuramos qual a probabilidade de isto acontecer. A Simulação cuidou melhor esta situação, porém na Árvore de Decisão foi possível fazer uma análise de sensibilidade em relação a esse parâmetro arbitrado, o que diminui bastante a desvantagem.

Outra vantagem para a Simulação é o fato de a Árvore de Decisão precisar considerar uma única variável como sendo o objetivo (no caso, o lucro a ser maximizado). Na Simulação, é possível verificar o desempenho de cada uma das 31 possíveis soluções nos critérios que forem considerados necessários – no caso, além do lucro, o risco envolvido na solução e a própria quantidade de propostas elaboradas. Estes últimos foram usados como critérios de desempate, possibilitando uma análise mais rica e, assim, uma decisão mais embasada.

Vantagens e Desvantagens das Metodologias

No [Quadro 4.4](#) apresentamos um resumo das vantagens e desvantagens detalhadas anteriormente.

CONSIDERAÇÕES FINAIS

Este capítulo apresentou um problema real de determinação de mix de projetos de consultoria, em que não se sabia, *a priori*, quais das propostas elaboradas se transformariam efetivamente em projetos a serem feitos. Tal presença de incerteza é suficiente para inviabilizar a abordagem mais tradicional para este tipo de problema: a Programação Binária.

Tratou-se tal problema de forma não probabilística pelos critérios Maximax, Maximin e Minimax, e de forma probabilística a partir de uma Árvore de Decisões e da análise de risco viabilizada pela Simulação, que se revelou adequada para tratar tal tipo de problema, sobretudo pelo fato de permitir que riscos sejam incorporados

QUADRO 4.4 Quadro comparativo das metodologias

Não probabilísticas		
	Vantagem	Desvantagem
	Não há dependência de subjetividade presente nas probabilidades	Todas as propostas são tratadas da mesma maneira
Critérios		
Maximax		Apresentou múltiplas soluções em todos os cenários
Maximin		Não funcionou para este problema
Minimax	Apresentou apenas uma solução diária para cada cenário	
	Sensibilidade em relação à multa	
Probabilísticas		
	Vantagem	Desvantagem
	Cada proposta é tratada de forma diferente	Há dependência de subjetividade
Abordagens		
Árvore da Decisão	Apresentou uma solução ótima	Confecção muito trabalhosa
	Contemplou o perfil do risco do tomador de decisão	Foi arbitrada uma multa para penalizar o lucro
		Foi escolhida apenas uma variável objetivo
Simulação	Foi verificado o desempenho de cada uma das 31 soluções por vários critérios	Planilha muito “pesada”
	Foi verificada a probabilidade da demanda por recursos superar a oferta	Não apresentou diretamente uma solução ótima
		Não contemplou o perfil de risco do tomador de decisão

de forma satisfatória, garantindo maior representatividade para o modelo e mais confiabilidade para os resultados.

Por causa da sua natureza, a Simulação apresenta uma desvantagem: a solução ótima não é revelada. Neste caso em particular, isso não representou um grande problema, pois a quantidade de soluções possíveis era finita e razoavelmente pequena, e foi possível verificar o desempenho de cada uma delas. Isso revelou outra vantagem da ferramenta: a possibilidade de considerar, para a escolha da solução, não apenas uma variável-objetivo (o lucro esperado, no caso), mas também outros aspectos importantes, como o risco de desgaste comercial e a quantidade de propostas elaboradas, referentes a cada uma das possíveis soluções.

EXERCÍCIOS

1. Diariamente, um jornaleiro compra Q exemplares de um jornal, ao custo unitário de R\$ 1,00. Cada jornal é vendido por R\$ 3,00, sendo as sobras ao final do dia perdidas (valor residual nulo). A demanda diária de jornais, que é independente e identicamente distribuída para todos os dias, é descrita pela seguinte distribuição de probabilidades:

Demanda diária (D)	Probabilidade (%)
60	16
70	19
80	24
90	22
100	19

O objetivo do problema é determinar o valor ótimo de Q (a quantidade diária de jornais comprada pelo jornaleiro).

- a. Trate o problema de forma determinística. Para tal, considere a demanda diária como um parâmetro fixo, arredondando o valor esperado da sua distribuição.
- b. Agora trate o problema de forma não determinística; mas ainda não considere as probabilidades, apenas a incerteza de a demanda poder assumir qualquer dos valores apresentados na tabela. Resolva utilizando cada um dos critérios não probabilísticos.
- c. Inclua agora as probabilidades e resolva o problema através de uma Árvore de Decisão.
- d. Ainda abordando de forma probabilística, simule, para os valores razoáveis de Q (60, 70, 80, 90 e 100), a atividade do jornaleiro por um período de 100 dias. Para cada dia simulado, liste a demanda observada, a quantidade vendida, a receita, o custo e o lucro diário. Ao final da corrida, calcule a média para cada uma destas variáveis e o desvio-padrão do lucro. Faça um gráfico da média e do desvio-padrão do lucro diário em função de Q .
- e. Reflita sobre as vantagens e desvantagens de cada abordagem, assim como sobre a sua adequação a este e a outros problemas.

REFERÊNCIAS

- CAMPOLINA, A. G.; CICONELLI, R. M. Qualidade de vida e medidas de utilidade: parâmetros clínicos para as tomadas de decisão em saúde. *PAN AMERICAN JOURNAL OF PUBLIC HEALTH*. Volume 19, 2, p. 128-36, March 15, 2006.
- CORRAR, L. J.; THEÓFILO, C. R. (coordenadores). *Pesquisa operacional para decisão em contabilidade e administração*. São Paulo: Atlas. 2004.
- FERREIRA, J. S. Decisão, 2004. Disponível em: http://www2.egi.ua.pt/cursos_2004/files/SAD/Decisão.pdf Acesso em 15/06/2009.
- GOMES, L. F.; GOMES, C. F.; ALMEIDA, A. T. *Tomada de decisão gerencial – Enfoque multicritério*. 3. ed. São Paulo: Atlas. 2009.
- HERTZ, D. B. Análise de risco em investimentos de capital. *Biblioteca Harvard*. 1980.
- HILLIER, F. S. *Introdução à pesquisa operacional*. 8. ed. São Paulo: McGraw Hill. 2006.
- LEÃO, F. B.; SILVA, L. G. W; MANTOVANI, J. R. S. Localização de faltas em sistemas de energia elétrica através de um modelo de programação binária e algoritmo genérico. *XVI CBA*, 2006.
- MEDEIROS, V. Z. *Métodos quantitativos com Excel*. São Paulo: Cengage Learning. 2008.
- MENDES, E. L.; SOARES, T. M.; SOUZA, R. C. Escores de variáveis latentes: Uma opção para o índice ANEEL de satisfação do consumidor. *APDIO – Associação Portuguesa de Investigação Operacional*, 26, p. 211-25, 2006.
- MOORE, J. H.; WEATHERFORD, L. R. *Tomada de decisão em administração com planilhas eletrônicas*. 6. ed. Porto Alegre: Bookman. 2005.
- MOREIRA, D. A. *Pesquisa operacional: Curso introdutório*. São Paulo: Thomson Learning. 2007.
- VATTER, P. A.; BRADLEY, S. P.; JACKSON, B. B.; FREY, JR., S. C. *Quantitative methods in management: text and cases*. Homewood: Irwin. 1978.

Métodos de previsão

Métodos de previsão podem ajudar as empresas a, entre outros objetivos:

- planejar e estimar melhor os valores dos investimentos a serem realizados, os estoques a serem criados, a capacidade de atendimento necessária para a prestação de um serviço, o tamanho da produção etc.;
- verificar o efeito da entrada de concorrentes;
- descobrir onde é necessário o esforço de venda e se criar cronogramas diferenciados de promoções e descontos.

Como resultado do processo de previsão, estimativas superestimadas (previsão acima do real) acarretam o chamado custo de excesso, que pode se manifestar na forma de, por exemplo, custo fixo maior, custo desnecessário de estoque, obsolescência, perecibilidade ou gasto maior de propaganda.

Já estimativas subestimadas (previsão abaixo do real) ocasionam o chamado custo de falta, que pode assumir a forma de, por exemplo, perda da margem de contribuição de um produto ou serviço, ordens em espera na fila de produção ou consequências negativas de uma demanda não atendida.

Por isso, o objetivo consiste em realizar previsões que sejam o mais próximo possível da realidade, ou seja, criar modelos que apresentem menores erros de previsão possíveis.

Existem métodos qualitativos e quantitativos de previsão. Os métodos quantitativos, que são o objeto desta Parte II, podem ser subdivididos em dois tipos: os causais e os de série temporal.

Os métodos causais são recomendados quando a variável a ser predita está fortemente correlacionada com outras variáveis que estão sujeitas a grandes variações. Eles costumam apresentar resultados bastante satisfatórios, mas carregam, intrinsecamente, dois inconvenientes:

1. é necessário o acesso ao histórico não só da variável a ser prevista, como também das variáveis capazes de impactar a grandeza que está se querendo estimar;
2. é preciso saber prever o valor futuro destas variáveis potencialmente impactantes, chamadas variáveis explicativas, para que a grandeza em questão possa ser bem estimada.

O representante clássico dos métodos causais é a Regressão Múltipla, abordada nos Capítulos 5 e 6. Mediante a Regressão Múltipla, tenta-se prever, em um *call center* que recebe chamadas de clientes com reclamações ou dúvidas acerca de sua conta telefônica, respectivamente, o volume e a duração média de ligações recebidas diariamente, em função de diversos fatores, como o dia da semana, a ocorrência de feriados e a data da chegada e do vencimento da conta.

Já os métodos de série temporal¹ pressupõem que os fatores temporais que influenciaram padrões da atividade no passado continuarão a fazê-lo, de forma similar, no futuro. Em vista disso, eles envolvem a projeção de valores futuros de uma variável, com base, *inteiramente*, em observações do histórico dessa variável; em outras palavras, outras variáveis não entram na análise e, naturalmente, não precisam ter seus valores coletados nem previstos.

Portanto, esses métodos acabam sendo mais simples e, o que é mais importante, podem ser aplicados em uma gama maior de situações,² quando comparados aos modelos causais. No entanto, algumas vezes, seus resultados não costumam ser tão satisfatórios.

Existem diversos métodos de séries temporais,³ sendo um dos mais completos o da Decomposição Clássica, que tem, como consequência natural, a sua complexidade, que pode ser considerada relativamente alta, sobretudo no que diz respeito aos procedimentos operacionais. Em função disso, este livro oferece, como último material desta Parte II, um tutorial multimídia (texto-planilha) sobre o método da Decomposição Clássica, que está apresentado no Apêndice.

1. Uma série temporal consiste em um conjunto de dados numéricos obtidos durante períodos regulares ao longo do tempo.

2. Notadamente quando não estão disponíveis informações acerca de outras variáveis.

3. Média Móvel, Amortecimento Exponencial Simples, Análise de Tendência, Método de Holt, Modelo de Winter etc.

Previsão da demanda de ligações em um *call center* por meio de um modelo de Regressão Múltipla

Marco Aurélio Carino Bouzada
MADE/UNESA

Objetivos de aprendizagem do capítulo

- Saber construir um modelo causal, capaz de explicar o comportamento de uma grandeza real com base no comportamento de outras variáveis.
- Entender as vantagens de um modelo causal em relação a um modelo de série temporal.
- Aprender a quantificar variáveis qualitativas e incluí-las no modelo por meio de variáveis *dummy*.
- Conseguir aferir a capacidade de explicação do modelo.
- Conseguir aferir a capacidade de explicação de cada variável e entender como cada uma impacta a variável dependente.
- Saber refinar o modelo, tornando-o mais parcimonioso com a retirada de variáveis pouco explicativas.
- Entender como obter e interpretar a equação capaz de prever a variável dependente.
- Aprender a mensurar o erro de previsão decorrente da utilização de modelo desenvolvido.

INTRODUÇÃO

Call centers consistem em centros operacionais, instalados para o uso da telecomunicação e das tecnologias computacionais, para automatizar uma quantidade volumosa de diferentes atividades e serviços telefônicos, tanto recebidos quanto originados pelo centro. Os centros do tipo *inbound*, onde as chamadas são originadas pelos clientes, são caracterizados como um sistema constituído de várias pessoas (atendentes), que recebem ligações telefônicas de outras pessoas, normalmente clientes – ainda que potenciais, que desejam obter alguma informação, contratar algum serviço, comprar

algum produto, responder a alguma pesquisa, atualizar dados, registrar ocorrências ou fazer reclamações, entre outras demandas.¹

Os gerentes e planejadores dos *call centers* exercem funções que, hoje em dia, exigem muito mais empenho e capacidade do que no passado. Com mais produtos e serviços sendo criados, disponibilizados no mercado, vendidos e assistidos, em relação a momentos anteriores, eles precisam entender o que está havendo nos *call centers* para saber como as suas características influenciam os indicadores de desempenho. Certamente a utilização de simples palpites, experiências de tentativa e erro, intuição ou softwares de funcionamento invisível para o usuário não representam a melhor alternativa para os gerentes das centrais de empresas que querem ser bem-sucedidas.²

Os supervisores dos *call centers*, responsáveis pelo planejamento operacional, precisam dimensionar eficientemente as suas equipes de modo a proporcionar níveis de serviço satisfatórios a custos razoáveis. A administração apropriada das centrais requer a estimação de algumas características operacionais que levam ao dimensionamento recomendado da equipe de atendimento; dentre essas características, destaca-se a demanda de ligações de clientes, questão a ser permanentemente abordada neste tipo de negócio.³

Considerando-se que o custo de pessoal representa aproximadamente 70% do total da indústria, a matéria-prima para o processo de gerenciamento eficiente (níveis satisfatórios de serviço a custos razoáveis) de um *call center* consiste na previsão da demanda de ligações e da consequente carga de trabalho, que precisam, então, ser estimadas da forma mais acurada possível. Essa necessidade confere à previsão de demanda de ligações o status de atividade crítica durante a busca pela eficiência e pela efetividade do negócio. Uma previsão acurada dos parâmetros do sistema é um pré-requisito para um nível de serviço consistente, representando um dos passos mais importantes do processo como um todo.⁴

A previsão do volume de chamadas é um importante aspecto das operações em *call centers*, sendo necessária para a previsão de carga de trabalho, alocação de mão de obra e planejamento de capacidade. Em geral, existem algumas regularidades nos dados de chegada das chamadas, como a dependência entre os dias, ou seja, a correlação entre os volumes de dias consecutivos e possíveis ciclos semanais, mensais e até anuais nos volumes de ligações; e a dependência intradiária, ou seja, a correlação entre chegadas em diferentes períodos de tempo – manhã, tarde ou noite – dentro do mesmo dia.⁵

O processo de previsão consiste tanto em arte quanto em ciência. Em arte, porque o futuro, afinal de contas, está sendo previsto e a acurácia da previsão vai depender em parte do julgamento e da experiência de quem estiver conduzindo o processo. Mas também em ciência: um processo matemático, passo a passo, que utiliza a história passada para prever eventos futuros. Assim, um entendimento de técnicas estatísticas especializadas é necessário para o processo ser benfeito. Os administradores que

1. Grossman *et al.* (2001); Hawkins *et al.* (2001).

2. Mehrotra; Profozich; Bapat (1997).

3. Weinberg; Brown; Stroud (2006); Klungle; Maluchnik (1997).

4. Hall; Anton (1998); Alam (2002); Gans; Koole; Mandelbaum (2003); Reynolds (2005); Bouzada (2006); Weinberg; Brown; Stroud (2006).

5. Shen; Huang (2005).

possuem softwares de gerenciamento de força de trabalho, que automatizam o processo de previsão, não podem pensar que isso apenas é suficiente; entender estes cálculos é tão crítico quanto possuir o software: não só para verificar a acurácia dos resultados como também, e talvez de forma mais importante, para explicar os números para a gerência. Ou seja, mesmo tendo as ferramentas, vale a pena entender os fundamentos do processo de previsão.⁶

Para gerar previsões acuradas dos volumes futuros de chamadas, os gerentes dos *call centers* precisam juntar informações de várias fontes, desde previsões do mercado até, eventualmente, a duração dos intervalos comerciais veiculados na mídia. No entanto, parte importante das informações provém das estatísticas do histórico de chamadas que, usualmente, podem ser ajustadas para as tendências atuais ou esperadas.⁷

Os modelos de séries temporais têm sido a abordagem mais utilizada para o tratamento da previsão de chamadas telefônicas. No entanto, outros modelos de previsão foram aplicados com sucesso em alguns ambientes de *call center*, como, por exemplo, centrais de vendas cuja previsão de chamadas se baseia nas remessas de produtos recebidas e nas propagandas enviadas aos potenciais clientes.⁸

A atividade de previsão é inerentemente difícil no caso dos *call centers* em razão do tamanho relativamente pequeno dos intervalos de tempo utilizados na prática. Técnicas de previsão estabelecidas, como o Amortecimento Exponencial de Winters e a Regressão Linear Múltipla, são úteis para este tipo de negócio. Geralmente, usar uma abordagem de previsão apropriada irá reduzir pela metade os erros de previsão. No entanto, muitos *call centers* encontram dificuldades nesse processo por conta do conhecimento técnico necessário para adequar estas técnicas aos complexos padrões de chamadas e ao aspecto pouco organizado dos dados.⁹

Em geral, se a variável a ser predita está fortemente correlacionada com outras variáveis que estão sujeitas a grandes variações, um modelo multivariado é necessário. O método da Regressão Múltipla, em particular, é relativamente robusto em relação a afastamentos dos pressupostos da normalidade e da homocedasticidade. Além disso, a eventual ocorrência de multicolinearidade entre as variáveis explicativas não prejudica a capacidade preditiva do modelo, causando apenas problemas quanto à sua parcimônia e de compreensão do real efeito das variáveis independentes sobre o comportamento da variável dependente.¹⁰

O PROCESSO DE PREVISÃO DE DEMANDA NO CALL CENTER ESTUDADO SEM A REGRESSÃO MÚLTIPLA

A principal ferramenta que a empresa estudada utiliza para auxiliar as decisões gerenciais de planejamento operacional é um software totalmente integrado à infraestrutura de telefonia. O produto consiste em uma ferramenta de força de trabalho utilizada

6. Reynolds (2005); Bouzada (2006).

7. Matan; Nourbakhsh (1998).

8. Mehrotra (1997).

9. Grossman *et al.* (2001).

10. Silver (2000); Levine; Berenson; Stephan (2000); Corrar; Paulo; Dias Filho (2007).

para dimensionar e controlar o atendimento, que auxilia os gerentes – entre outras atividades – na previsão de demanda (por horário e dia da semana).

De acordo com a gerência da empresa, a principal preocupação consiste em conseguir melhorar as previsões de demanda de ligações. A acurácia das previsões não é boa para alguns produtos. Para o produto que abrange clientes que buscam serviços relacionados à telefonia fixa, como mudança de endereço de cobrança e dúvidas ou problemas com a conta telefônica, a empresa dispõe de todas as informações que poderiam fornecer uma melhor previsão, estando apenas carente de um método mais adequado. Essas informações são basicamente as quantidades de contas enviadas para os clientes em cada data, normalmente cinco a seis dias antes do vencimento, e as suas próprias datas de vencimento. A impressão da gerência é que a demanda de ligações para o produto é influenciada pelos eventos de chegada da conta e pelo seu próprio vencimento, datas em que haveria maior acesso ao serviço por parte dos clientes. O problema é que o sistema utilizado para previsão de demanda não leva em conta essas informações.

A magnitude do erro absoluto percentual médio, segundo a ótica diária, incorrido durante o processo de previsão de demanda para o produto em questão era da ordem de 5%. Na opinião dos responsáveis pela previsão de demanda, esse indicador de erro encontra-se em um patamar baixo, sendo considerado mais do que aceitável. No entanto, estes responsáveis entendem que existe um potencial de melhoria no processo, de forma que esses erros podem se tornar ainda menores.

O primeiro passo do processo de previsão de demanda para o produto em questão consiste na geração de uma previsão básica, realizada pelo software de gerenciamento de força de trabalho. O segundo passo consiste em uma crítica, por parte dos analistas da coordenação, a esses números gerados pelo software. O terceiro passo consiste em incorporar aos valores previstos o possível impacto causado por alguns eventos específicos, como a segunda-feira de Carnaval ou uma ocorrência diferente em algum sábado, por exemplo. A maneira de incorporar tais impactos depende de o evento ser novo ou desconhecido e do prévio conhecimento das suas consequências.

Alguns dos eventos especiais podem ter seu impacto previsto de uma forma melhor, quando a equipe de analistas é subsidiada por algumas premissas recebidas de outros setores capazes de impactar a demanda de chamadas, como: previsão de crescimento da base de clientes; implantação de novos planos e serviços na telefonia fixa, como o plano de minutos, por exemplo.

Essas premissas também envolvem o comportamento dos ciclos das contas telefônicas dos clientes, uma informação oriunda do setor de faturamento da empresa telefônica e que ajuda a previsão de demanda do produto em questão. A informação consiste, basicamente, nas quantidades de contas enviadas para os clientes em cada data, normalmente cinco a seis dias antes do vencimento. Ela é útil porque a demanda de ligações para o produto é influenciada, segundo a equipe de previsão, pelos eventos de chegada da conta no destino e seu vencimento, datas em que haveria maior acesso ao serviço por parte dos clientes, para fins de esclarecimento e/ou reclamação.

Na verdade, durante todo o período de quatro a cinco dias em que a conta está com o cliente (desde a chegada até o vencimento), ocorre um acréscimo no volume de ligações, mas o maior impacto acontece no momento da chegada da conta na

residência do cliente, na opinião da coordenação responsável pelo produto. Segundo os analistas, o impacto pós-vencimento existe quando o cliente que esqueceu de pagar a conta liga para saber como proceder, mas é pequeno, não sendo significativo em termos de aumento de demanda.

O problema é que o software não leva em conta isso; a sazonalidade do dia da semana é considerada, mas não a distância do dia em relação à chegada da conta ou ao seu vencimento. Por exemplo, a conta de determinado cliente vence no dia 10 de cada mês, e não na segunda quarta-feira; então, no mês em que o dia 10 caísse em uma terça-feira, esse dia da semana deveria ser mais impactado do que os outros, enquanto no mês em que o dia 10 caísse em uma quarta-feira, o impacto maior deveria ser incluído neste dia. Mas o software considera, em termos de sazonalidade, todas as terças-feiras da mesma forma, seguindo o mesmo raciocínio para os outros dias da semana. Além disso, há meses em que as contas não são enviadas exatamente cinco dias antes do vencimento, e a chegada da conta ao destino também deve ser considerada como um evento especial.

Essas informações são levadas em conta no processo de previsão, mas não de forma automática. Toda a análise é feita pela equipe de previsão, e não pelo software. Mas durante esse processo, os analistas não levam em conta um possível impacto da véspera e da antevéspera do vencimento da conta e nem do período pós-vencimento; apenas o efeito dos eventos chegada e vencimento.

O PROCESSO DE PREVISÃO DE DEMANDA COM A REGRESSÃO MÚLTIPLA

A ideia consiste em fugir dessa pós-análise não metódica – que depende muito da intuição do analista – e poder contar com uma ferramenta que use as informações do faturamento no processo de elaboração das previsões, para que se possa evitar o trabalho manual, que aumenta a possibilidade de erro. A ferramenta não tem a pretensão de substituir a intuição e a experiência da equipe de previsão, mas servir de apoio à tomada de decisão. Ou seja, a opinião dos especialistas não deve ser descartada; ela deve, sim, ser utilizada para enriquecer o modelo que busca, basicamente, objetivizar as informações subjetivas que são relevantes para o processo de previsão.

O método escolhido foi o da Regressão Múltipla, pois é capaz de captar o impacto específico de cada evento e informação na demanda de cada dia. Por exemplo, qual o impacto, na demanda de ligações, do fato de o dia em questão cair em uma quinta-feira? Qual o impacto de o dia consistir em um feriado? E de chegarem 200 mil contas para os clientes no dia em questão? E de o dia ser a véspera do vencimento de 500 mil contas?

A ideia é verificar a distância do dia específico em relação aos eventos especiais, a chegada e o vencimento das contas, e calcular para quantas contas esse é o dia C (dia de chegada), o V (dia do vencimento), o $V - 1$ (véspera do vencimento), o $V - 2$ (antevéspera do vencimento), ..., $V + 1$ (dia seguinte ao vencimento), $V + 2$, ... entre outros.

De posse do cronograma de faturamento fornecido pela equipe de previsão, para o período de julho de 2005 a julho de 2006, foi possível consolidar as informações em termos da data de chegada e da data de vencimento das contas, conforme apresentado parcialmente nas Tabelas 5.1 e 5.2, respectivamente.

TABELA 5.1 Contas agrupadas por data de chegada prevista (setembro a outubro de 2005)

Data de Chegada Prevista	Total
06/09/05	3.360.569
13/09/05	801.143
14/09/05	432.670
15/09/05	709.437
26/09/05	1.315.342
27/09/05	820.855
28/09/05	1.555.097
29/09/05	2.506.038
04/10/05	1.808.735
05/10/05	1.167.447
07/10/05	800.105
11/10/05	1.890.364
18/10/05	709.437
26/10/05	1.731.518
28/10/05	587.118
Total	20.195.875

Para poder ser desenvolvido o modelo de regressão múltipla para prever a quantidade de ligações em cada data, foi preciso ter acesso ao histórico desta grandeza. Por motivo de sigilo estratégico, esses valores não puderam ser revelados, em sua plena magnitude; no entanto, a eles foi aplicada uma constante de proporcionalidade, e os resultados, de julho de 2005 até março de 2006, fornecidos pela empresa, estão apresentados graficamente na [Figura 5.1](#).

Como se observa, a série não apresenta tendência de crescimento ou de decréscimo, mas apenas um comportamento errático, dispensando, assim, a necessidade de inclusão de uma variável associada à passagem do tempo.

De posse de todas essas informações foi possível montar uma planilha com os dados necessários para a construção do modelo de regressão múltipla dentro do qual o comportamento da variável dependente quantidade de ligações em um dia específico poderia estar sendo explicado – inicialmente – por 10 variáveis independentes:

- dia da semana;
- feriado – se o dia específico consiste em um feriado ou não;

TABELA 5.2 Contas agrupadas por data de vencimento (setembro a outubro de 2005)

Vencimento	Total
04/09/05	2.142.215
07/09/05	1.918.920
11/09/05	2.177.552
13/09/05	798.630
15/09/05	1.456.656
17/09/05	801.143
20/09/05	432.670
23/09/05	709.437
01/10/05	2.136.197
04/10/05	2.142.215
07/10/05	1.918.920
11/10/05	2.177.552
13/10/05	798.630
15/10/05	1.456.656
17/10/05	801.143
20/10/05	432.670
23/10/05	709.437
Total	23.010.643

- chegada (C) – quantas contas estão previstas para chegarem ao cliente no dia específico;
- vencimento (V) – quantas contas vencem no dia específico;
 - V – 3 – para quantas contas aquele dia específico corresponde a três dias antes do vencimento, ou, em outras palavras, quantas contas irão vencer três dias depois da data específica;
 - V – 2 – para quantas contas aquele dia específico corresponde à antevéspera do vencimento;
 - V – 1 – para quantas contas aquele dia específico corresponde à véspera do vencimento;
 - V + 1 – para quantas contas aquele dia específico corresponde a um dia depois do vencimento, ou, em outras palavras, quantas contas terão vencido na véspera daquela data específica;

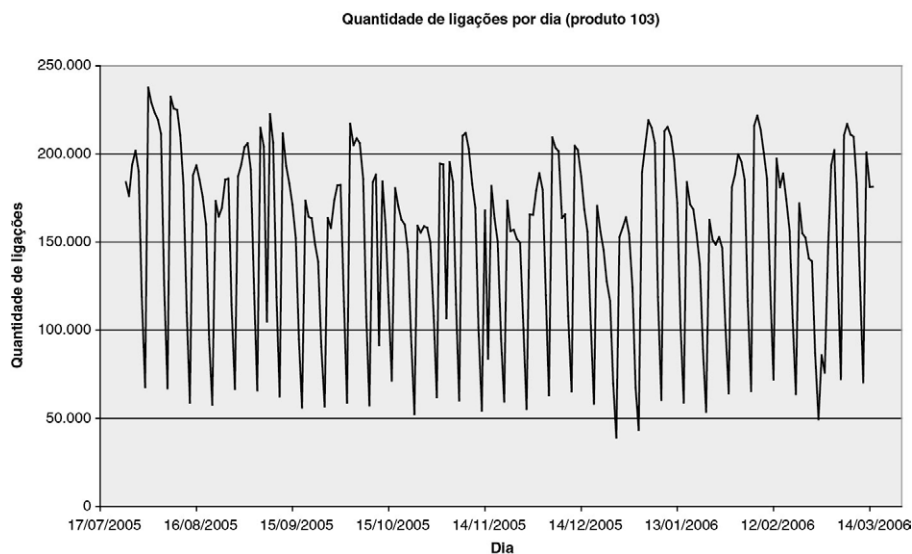


FIGURA 5.1 Quantidade de ligações por dia (julho de 2005 a março de 2006).

- $V + 2$ – para quantas contas aquele dia específico corresponde a dois dias depois do vencimento;
- $V + 3$ – para quantas contas aquele dia específico corresponde a três dias depois do vencimento.

A Tabela 5.3 apresenta, em parte, a planilha de dados em questão.

Cabe salientar que as variáveis independentes dia da semana e feriado foram trabalhadas como sendo do tipo *dummy* (0 = Não; 1 = Sim). Foi criada uma variável *dummy* para a ocorrência de feriado e seis variáveis *dummy* para o dia da semana: domingo, segunda-feira, terça-feira, quarta-feira, quinta-feira e sexta-feira.

Dessa forma, para o dia 28/02/2006, por exemplo, a variável *dummy* feriado assumiu o valor 1, por se tratar da terça-feira de Carnaval; as variáveis *dummy* domingo, segunda-feira, quarta-feira, quinta-feira e sexta-feira assumiram o valor 0, enquanto a variável *dummy* terça-feira assumiu o valor 1. Seguindo o mesmo raciocínio, para o dia 04/02/2006, por exemplo, a variável *dummy* feriado assumiu o valor 0; e as variáveis *dummy* domingo, segunda-feira, terça-feira, quarta-feira, quinta-feira e sexta-feira assumiram o valor 0, caracterizando, portanto e por exclusão, o dia em questão como sendo um sábado.

Após a regressão ser executada, a primeira análise aponta para um modelo com um potencial de explicação muito bom, já que o R-quadrado ajustado indica que 84% da variação dos dados pode ser explicada. Neste caso, pode-se inferir a real influência das variáveis explicativas na quantidade de ligações e aceitar o modelo com elevada confiança, já que o nível de significância da estatística F foi igual a 0,0000.

A contribuição de cada variável independente é expressa pelos coeficientes estimados, juntamente com os respectivos *p-values* (ou valores *p*), apresentados na Tabela 5.4.

TABELA 5.3 Planilha de dados do modelo de regressão múltipla para prever a quantidade de ligações (fevereiro de 2006)

Dia	Dia da semana	Feriado?	Chegada (C)	V – 3	V – 2	V – 1	Vencimento (V)	V + 1	V + 2	V + 3	Ligações
01/02/200	4ª-feira	Não	2.506.038	2.142.215	-	-	2.136.197	-	-	-	199.908
02/02/200	5ª-feira	Não	-	-	2142.215	-	-	2.136.197	-	-	195.497
03/02/200	6ª-feira	Não	1.122.892	-	-	2.142.215	-	-	2.136.197	-	185.542
04/02/200	Sábado	Não	-	1.918.920	-	-	2.142.215	-	-	2.136.197	116.657
05/02/200	Domingo	Não	-	-	1.918.920	-	-	2.142.215	-	-	65.558
06/02/200	2ª-feira	Não	587.354	-	-	1.918.920	-	-	2.142.215	-	216.104
07/02/200	3ª-feira	Não	1.484.473	-	-	-	1918.920	-	-	2.142.215	221.831
08/02/200	4ª-feira	Não	1.238.119	2.177.552	-	-	-	1.918.920	-	-	213.584
09/02/200	5ª-feira	Não	801.143	-	2.177.552	-	-	-	1.918.920	-	200.359
10/02/200	6ª-feira	Não	-	798.630	-	2.177.552	-	-	-	1.918.920	185.882
11/02/200	Sábado	Não	-	-	798.630	-	2.177.552	-	-	-	122.764
12/02/200	Domingo	Não	-	1.456.656	-	798.630	-	2.177.552	-	-	72.033
13/02/200	2ª-feira	Não	-	-	1.456.656	-	798.630	-	2.177.552	-	197.231
14/02/200	3ª-feira	Não	432.670	801.143	-	1.456.656	-	798.630	-	2.177.552	181.092
15/02/200	4ª-feira	Não	709.437	-	801.143	-	1.456.656	-	798.630	-	188.990
16/02/200	5ª-feira	Não	-	-	-	801.143	-	1.456.656	-	798.630	174.253
17/02/200	6ª-feira	Não	-	432.670	-	-	801.143	-	1.456.656	-	156.317

(Continua)

TABELA 5.3 Planilha de dados do modelo de regressão múltipla para prever a quantidade de ligações (fevereiro de 2006) (*cont.*)

Dia	Dia da semana	Feriado?	Chegada (C)	V – 3	V – 2	V – 1	Vencimento (V)	V + 1	V + 2	V + 3	Ligações
18/02/200	Sábado		Não.	-	432.670	-	-	801.143	-	1.456.656	105.306
19/02/200	Domingo	Não	-	-	-	432.670	-	-	801.143	-	63.827
20/02/200	2ª-feira	Não	-	709.437	-	-	432.670	.	-	801.143	171.921
21/02/200	3ª-feira	Não	-	-	709.437	-	-	432.670	-	-	155.035
22/02/200	4ª-feira	Não	2.136.197	-	-	709.437	-	-	432.670	-	152.537
23/02/200	5ª-feira	Não	1.555.097	-	-	-	709.437	-	-	432.670	140.621
24/02/200	6ª-feira	Não	2.506.038	-	-	-	-	709.437	-	-	139.056
25/02/200	Sábado	Não	-	-	-	-	-	-	709.437	-	86.746
26/02/200	Domingo	Não	-	2.136.197	-	-	.	-	-	709.437	49.520
27/02/200	2ª-feira	Sim	-	-	2.136.197	-	-	-	-	-	85.790
28/02/200	3ª-feira	Sim	-	-	-	2.136.197	-	-	-	-	75.801

TABELA 5.4 Coeficientes estimados das variáveis independentes do modelo de regressão múltipla para prever a quantidade de ligações na data em questão

	Coeficientes	valor-P
Interseção	92.545	0,00
Domingo	(41.228)	0,00%
2ª-feira	79.882	0,00%
3ª-feira	75.694	0,00%
4ª-feira	73.812	0,00%
5ª-feira	73.068	0,00%
6ª-feira	61.763	0,00%
Feriado	(71.088)	0,00%
Chegada (C)	0,0035	13,57%
V - 3	0,0052	2,21%
V - 2	0,0042	7,81%
V - 1	0,0068	0,47%
Vencimento (V)	0,0074	0,19%
V + 1	0,0071	0,30%
V + 2	0,0064	0,75%
V + 3	0,0009	68,36%

Na Tabela 5.4, observa-se um elevado valor p para a variável $V + 3$, indicando que a quantidade de contas para as quais aquele dia específico corresponde a três dias depois do vencimento não pode ser vista como um bom predictor para a quantidade de ligações na data em questão.

Assim sendo, o efeito pós-vencimento parece, então, acabar já no segundo dia após o vencimento da conta, já que os valores p de $V + 1$ e $V + 2$ são baixos; em outras palavras, os clientes ainda ligam um ou dois dias após o vencimento da conta, mas esse impacto desaparece a partir do terceiro dia, deixando de ser significativo.

A variável $V + 3$ pode estar contaminando o modelo e convém retirá-la da análise. Sem ela, um novo modelo de regressão foi gerado, continuando com um potencial muito bom e até ligeiramente melhor, já que 85% da variação dos dados pode ser explicada por ele.

Os coeficientes de regressão estimados, juntamente com os respectivos valores p , estão apresentados na Tabela 5.5.

Como pode ser observado, os valores p das outras variáveis mudaram muito pouco e, para este modelo modificado, apenas duas variáveis apresentam coeficientes angulares com valores p superiores a 5%: Chegada (C) e $V - 2$.

TABELA 5.5 Coeficientes estimados das variáveis independentes do modelo (sem V + 3) de regressão múltipla para prever a quantidade de ligações na data em questão

	Coeficientes	valor-P
Interseção	92.693	0,00%
Domingo	(41.103)	0,00%
2ª-feira	79.874	0,00%
3ª-feira	75.900	0,00%
4ª-feira	73.908	0,00%
5ª-feira	73.082	0,00%
6ª-feira	61.898	0,00%
Feriado	(71.345)	0,00%
Chegada (C)	0,0034	14,30%
V – 3	0,0052	2,11%
V – 2	0,0042	7,86%
V – 1	0,0072	0,13%
Vencimento (V)	0,0077	0,06%
V + 1	0,0071	0,30%
V + 2	0,0060	0,71%

Em relação a esta última variável, ela será mantida no modelo porque o valor p do seu coeficiente está abaixo de um valor também bastante utilizado como nível de significância nos testes de hipóteses (10%) e pelo fato de as variáveis V – 3 e V – 1 apresentarem-se também como bons previsores. Ora, não faria muito sentido manter no modelo V – 3, retirar V – 2 e manter V – 1, pois a continuidade estaria sendo quebrada; seria o mesmo que afirmar que os clientes ligam para o produto em questão três dias antes do vencimento e na sua véspera, mas não na antevéspera!

Já a variável Chegada (C) também será mantida no modelo, apesar de apresentar um valor p maior do que 10%, pelo fato de ter sido identificada, na opinião subjetiva e calcada na experiência da equipe de planejamento, como crítica e muito impactante na demanda de ligações, contrariando o resultado objetivo que indica um coeficiente não significativo para essa variável. Também é válido ressaltar que seu valor p (14,30%) não ficou tão acima de um nível de significância considerado bom (5%).

Essa intervenção subjetiva em um modelo tão objetivo traz à tona a já mencionada importância da utilização da opinião de especialistas na construção de um processo de previsão, mesmo nos casos em que esse é essencialmente sistemático.

Dessa forma, a retirada da variável $V + 3$ foi benéfica ao modelo, tornando-o mais parcimonioso e permitindo uma interpretação dos coeficientes obtidos.

O primeiro coeficiente corresponde ao valor da variável dependente quantidade de ligações quando todas as variáveis independentes assumirem o valor zero; ou seja, quando domingo, segunda-feira, terça-feira, quarta-feira, quinta-feira, sexta-feira, feriado, C , $V - 3$, $V - 2$, $V - 1$, V , $V + 1$ e $V + 2$ forem zero – ou, ainda, quando o dia for sábado, não feriado e não consistir da data de chegada, nem de nenhuma data dentro do intervalo de seis dias em torno do vencimento de alguma conta, a quantidade de ligações demandadas prevista é de 92.693.

Esse número pode ser encarado como uma quantidade básica, à qual devem ser adicionados os efeitos da ocorrência das variáveis independentes, quantificados pelos seus respectivos coeficientes angulares.

Assim, o fato de o dia ser domingo contribui com uma diminuição de 41.103 ligações previstas em relação ao número básico (92.693); a segunda-feira espera, segundo o modelo, 79.874 ligações a mais do que o dia básico (sábado); a terça-feira, 75.900 ligações a mais do que o sábado; a quarta-feira, 73.908 ligações a mais; a quinta-feira, 73.082 a mais; e a sexta-feira, 61.898 ligações a mais do que o dia básico.

Com essas informações, é possível elaborar um gráfico (Figura 5.2) que permite uma melhor visualização da sazonalidade do dia da semana.

A evolução da demanda de ligações ao longo da semana, considerando dias sem a ocorrência de feriados e não próximos da chegada ou do vencimento de contas, começa com um valor bem baixo no domingo, apresentando um grande aumento na segunda-feira; a demanda vai caindo suavemente ao longo da semana, tendo uma queda um pouco mais brusca na sexta-feira e bem mais intensa no sábado, antes de voltar ao patamar de domingo.

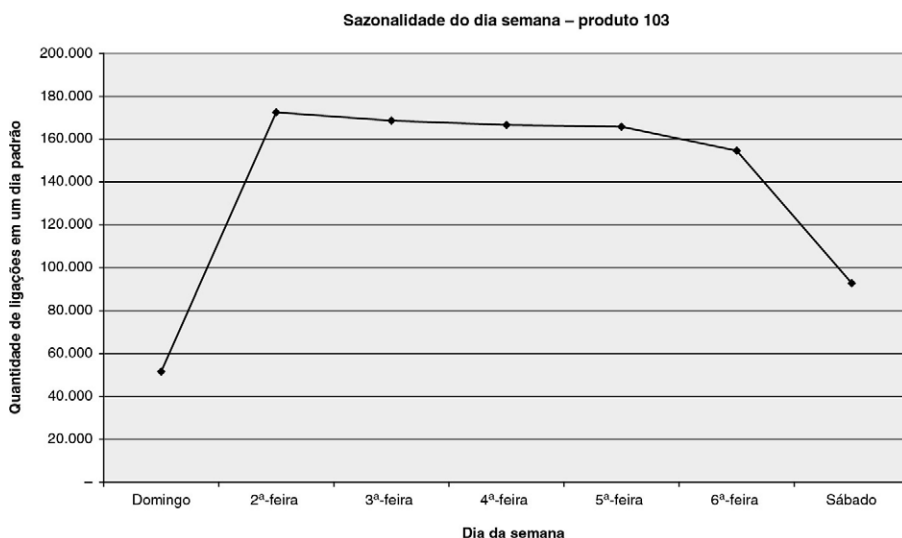


FIGURA 5.2 Comportamento sazonal do dia da semana – modelo sem $V + 3$.

De forma análoga aos dias da semana, o fato de um dia ser feriado reduz em 71.345 a quantidade prevista de ligações; e cada conta telefônica prevista para chegar à casa do cliente naquele dia aumenta em 0,0034 a previsão da quantidade de ligações demandadas para o dia em questão. Assim, se, por exemplo, 100 mil contas estiverem chegando em um dia específico, é de se esperar que isso implique um aumento de $100.000 \times 0,0034 = 340$ ligações no contingente diário. Se esse número for examinado de forma percentual, é possível concluir que, em média, 0,34% (= 0,0034) dos clientes ligam para a empresa telefônica no dia em que recebem a conta.

Em relação aos dias próximos à data do vencimento, é análogo o raciocínio que explica as consequências: cada conta vencendo três dias depois, ou seja, para as quais a data em questão representa $V - 3$, aumenta em 0,0052 a quantidade de ligações; cada conta vencendo dois dias depois aumenta em 0,0042 a quantidade de ligações; cada conta vencendo no dia seguinte aumenta em 0,0072 essa quantidade; cada conta vencendo no dia em questão aumenta em 0,0077 a mesma quantidade; cada conta vencendo na véspera aumenta em 0,0071 a quantidade; e cada conta vencendo na antevéspera aumenta em 0,0060 a quantidade prevista de ligações.

Naturalmente, os coeficientes não têm apenas significado isolado, mas também podem e devem ser analisados em conjunto: se as consequências dos efeitos dia da semana, ocorrência de feriado e data crítica em relação à chegada ou ao vencimento da conta forem incorporadas ao valor previsto para um dia básico, será possível estabelecer uma equação para prever a quantidade demandada de ligações em determinado dia em função das características da data em questão. Desta maneira, a quantidade de ligações pode ser prevista conforme a [Equação 5.1](#).

EQUAÇÃO 5.1 Equação do modelo para prever a quantidade de ligações

$$\begin{aligned} \text{Quantidade prevista de ligações} = & 92.693 - 41.103 \times \text{domingo} + \\ & 79.874 \times \text{segunda-feira} + 75.900 \times \text{terça-feira} + \\ & 73.908 \times \text{quarta-feira} + 73.082 \times \text{quinta-feira} + 61.898 \times \text{sexta-feira} \\ & - 71.345 \times \text{feriado} + 0,0034 \times C + 0,0052 \times V - 3 + \\ & 0,0042 \times V - 2 + 0,0072 \times V - 1 + \\ & 0,0077 \times V + 0,0071 \times V + 1 + 0,0060 \times V + 2 \end{aligned}$$

- as sete primeiras variáveis são do tipo *dummy*, assumindo o valor 0 (não) ou 1 (sim);
- as sete últimas representam, respectivamente: a quantidade de contas chegando no dia; vencendo três dias depois; dois dias depois; um dia depois; no próprio dia; um dia antes; e dois dias antes.

De posse dessa equação, é possível calcular quanto o modelo de regressão teria previsto para a quantidade de ligações em cada um dos dias acerca dos quais o histórico disponibiliza o valor real desta grandeza. Desta forma, pode-se comparar os valores que teriam sido previstos com os que efetivamente ocorreram, de maneira que um erro de previsão possa ser mensurado. Esta comparação pode ser visualizada na [Figura 5.3](#).

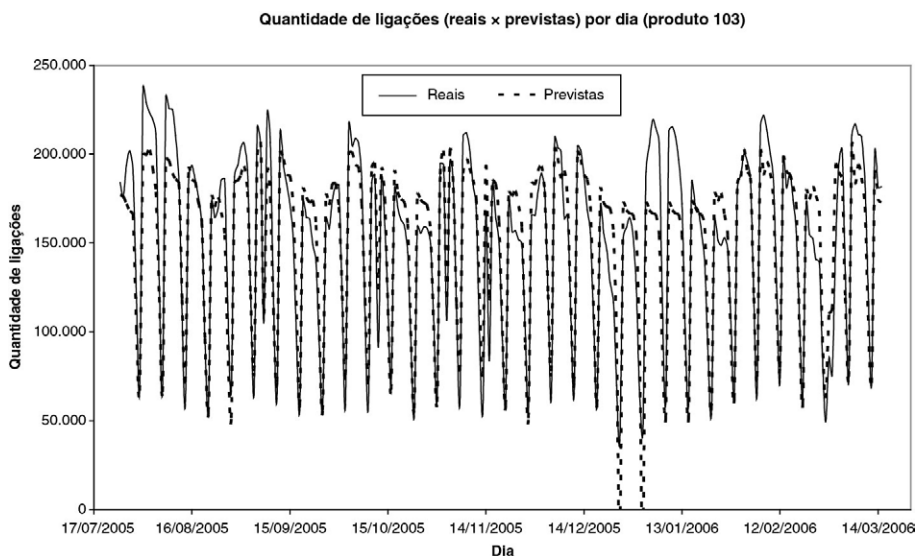


FIGURA 5.3 Valores reais multiplicados por uma constante e previstos pelo modelo para a quantidade de ligações por dia (julho de 2005 a março de 2006).

A diferença entre os valores reais e os gerados pelo modelo também é conhecida como resíduo. Para ser usado um modelo de regressão múltipla, é necessário supor que a correlação entre os resíduos, ao longo do espectro das variáveis independentes, é zero; ou seja, que os resíduos são independentes entre si, não apresentando, consequentemente, autocorrelação serial.¹¹

Uma forma de verificar a validade dessa suposição é por meio do teste de Durbin-Watson. No modelo que está sendo analisado, foi encontrado o valor 1,838 para a Estatística DW. Esta estatística obtida deve ser comparada com valores críticos que variam em função do nível de significância do teste, do tamanho da amostra e do número de variáveis independentes.¹²

Com 5% de significância, 234 dados na amostra e 14 variáveis independentes, o valor crítico inferior (d_L) obtido foi 1,661 e o valor crítico superior (d_U) encontrado foi 1,914.¹³

Como o valor obtido para a Estatística DW (1,838) se situa entre os dois valores críticos encontrados, o teste para verificar a existência de autocorrelação dos resíduos é não conclusivo: não é possível garantir que os resíduos são independentes, mas também é incorreto afirmar que existe uma relação de dependência entre eles.¹⁴

Sendo assim, parece válido examinar o gráfico do resíduo contra cada variável independente a fim de procurar uma suposta autocorrelação e a identificação de uma

11. Corrar; Paulo; Dias Filho (2007).

12. Corrar; Paulo; Dias Filho (2007).

13. Valores encontrados em tabela disponível no endereço <<http://www.stanford.edu/~clint/bench/dwcrit.htm>>. Acesso em 24/01/2009.

14. Corrar; Paulo; Dias Filho (2007).

eventual lei de formação nos valores dos resíduos à medida que o valor da variável independente avança. A Figura 5.4 ilustra o caso da variável *dummy* quarta-feira.

Como pode ser observado, não parece existir uma lei de formação; ou seja, a magnitude dos resíduos não parece depender do valor da variável independente (quarta-feira).

Os resíduos apresentaram comportamento semelhante em relação às outras variáveis *dummy* (feriado e as referentes aos outros dias da semana). E visto que o modelo trabalha com 14 variáveis independentes, seria um tanto inviável apresentar aqui tal quantidade de gráficos. Dessa forma, serão apresentadas graficamente apenas as relações entre os resíduos e as outras variáveis independentes mais marcantes: Chegada e Vencimento. Tais relações podem ser visualizadas nas Figuras 5.5 e 5.6, respectivamente.

O comportamento dos resíduos em relação às outras variáveis ($V - 3$, $V - 2$, $V - 1$, $V + 1$ e $V + 2$) é bastante similar ao verificado nas Figuras 5.5 e 5.6: parecem estar distribuídos aleatoriamente em torno do eixo das abscissas, não apresentando nenhuma tendência.

Tudo isso gera uma segurança maior para os resultados do modelo, já que o teste para a verificação de autocorrelação serial dos resíduos foi não conclusivo. É claro que esta última análise é meramente visual, carecendo de um cientificismo maior. Mas é importante proceder dessa maneira principalmente em situações em que os critérios objetivos apontam para uma indefinição, como foi o caso.

Já em relação à qualidade preditiva do modelo, foi possível perceber, na Figura 5.3, a menos de alguns pontos específicos, uma boa aderência visual dos valores previstos em relação aos dados reais. Esta aderência foi aqui formalizada pelo MAPE (Mean Absolute Percentual Error), ou erro médio percentual absoluto.



FIGURA 5.4 Resíduos do modelo X variável *dummy* quarta-feira.

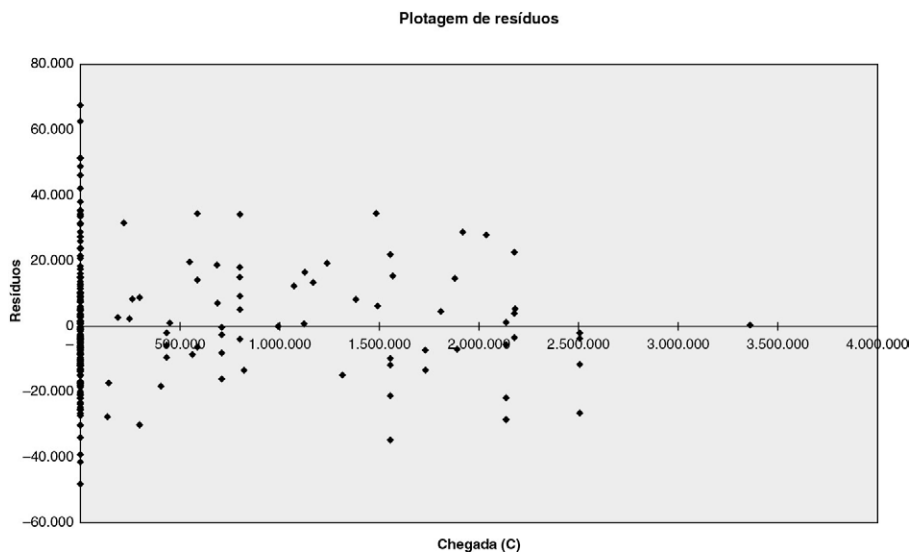


FIGURA 5.5 Resíduos do modelo X variável Chegada.

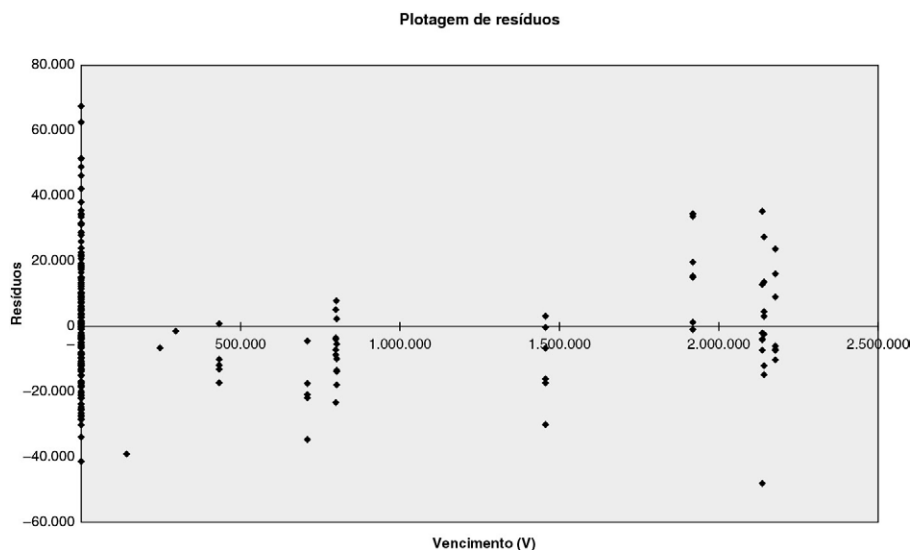


FIGURA 5.6 Resíduos do modelo X variável Vencimento.

A título de ilustração, a [Tabela 5.6](#) apresenta dados relativos ao mês de dezembro de 2005, nos quais constam a quantidade real de ligações, os valores que o modelo de regressão teria previsto para esta grandeza, assim como os erros percentuais absolutos oriundos desta previsão. Pode-se calcular que o modelo teria errado em 15,5%, em média, na previsão dos volumes diários de ligações durante o mês em questão.

TABELA 5.6 Valores reais, previstos e erros do modelo de regressão múltipla para prever a quantidade de ligações (dezembro de 2005)

Dia	Dia da semana	Feriado?	Ligações	Ligações (previstas pelo modelo)	Erro de previsão do modelo
01/12/2005	5ª-feira	Não	189.179	193.288	2,2%
02/12/2005	6ª-feira	Não	179.607	178.716	0,5%
03/12/2005	Sábado	Não	119.738	121.003	11%
04/12/2005	Domingo	Não	63.118	77.992	23,6%
05/12/2005	2ª-feira	Não	209.398	203.208	3,0%
06/12/2005	3ª-feira	Não	203.525	197.677	2,9%
07/12/2005	4ª-feira	Não	201.497	186.681	7,4%
08/12/2005	5ª-feira	Não	163.731	190.634	164%
09/12/2005	6ª-feira	Não	165.708	175.358	5,8%
10/12/2005	Sábado	Não	108.033	112.476	41%
11/12/2005	Domingo	Não	65.321	71.680	97%
12/12/2005	2ª-feira	Não	204.689	201.258	1,7%
13/12/2005	3ª-feira	Não	202.289	196.754	2,7%
14/12/2005	4ª-feira	Não	187.186	188.341	0,6%
15/12/2005	5ª-feira	Não	168.579	185.651	10,1%
16/12/2005	6ª-feira	Não	155.912	170.652	95%
17/12/2005	Sábado	Não	102.895	109.900	68%
18/12/2005	Domingo	Não	58.292	59.078	1,3%
19/12/2005	2ª-feira	Não	170.636	180.520	5,8%
20/12/2005	3ª-feira	Não	155.824	172.655	10,8%
21/12/2005	4ª-feira	Não	145.547	170.260	170%
22/12/2005	5ª-feira	Não	127.727	169.413	32,6%
23/12/2005	6ª-feira	Não	116.546	155.684	33,6%
24/12/2005	Sábado	Não	69.839	93.699	342%
25/12/2005	Domingo	Sim	39.038	118.895)	1484%
26/12/2005	2ª-feira	Não	152.909	172.567	12,9%
27/12/2005	3ª-feira	Não	158.507	168.593	64%
28/12/2005	4ª-feira	Não	164.109	166.601	1,5%
29/12/2005	5ª-feira	Não	154.690	165.775	7,2%
30/12/2005	6ª-feira	Não	124.406	154.591	24,3%
31/12/2005	Sábado	Não	67.541	92.693	37,2%

Mas dezembro é um mês atípico, com um comportamento geralmente mais difícil de prever. Tal afirmação pode ser verificada neste exemplo particular a partir da **Figura 5.7**, no qual é possível perceber que todos os outros meses relacionados, com exceção de janeiro, outro período atípico, apresentam erros médios de previsão pelo menos quatro pontos percentuais menores do que o erro médio de dezembro de 2005.

Em média, o modelo está incorrendo num erro de previsão da ordem de 10,98%. Embora não possa ser considerado um valor elevado, ele é substancialmente maior que os 5% em que a equipe de planejamento costuma incorrer na previsão de demanda do produto em questão. No entanto, cabem aqui algumas observações.

É válido lembrar, por exemplo, que, durante o processo padrão de previsão, os analistas recebem os valores previstos pelo software e acrescentam a eles os impactos esperados por eventos e premissas, como o comportamento do ciclo de contas e a ocorrência de feriados, assim como outros eventos especiais relacionados ao produto, como plano de minutos, campanhas na mídia, crescimento de base e outros potenciais aspectos impactantes. Os 5% de erro de previsão só são conseguidos após a consideração de todas essas subjetividades e dentro de um processo diário de revisão da previsão.

Ora, o modelo de regressão proposto visa substituir apenas os valores sugeridos pelo software, acrescidos da influência dos feriados e do comportamento dos ciclos de contas; nunca foi a ideia substituir o papel do analista encarregado da previsão – dotado de toda a sua valiosa experiência e subjetividade –, mas, sim, de fornecer-lhe um subsídio complementar para a tomada de decisão. Assim, e visto que o analista estará apto a acrescentar aos valores sugeridos pelo modelo de regressão os impactos que julgar relevantes, é mais justo comparar os erros de previsão do modelo aqui proposto

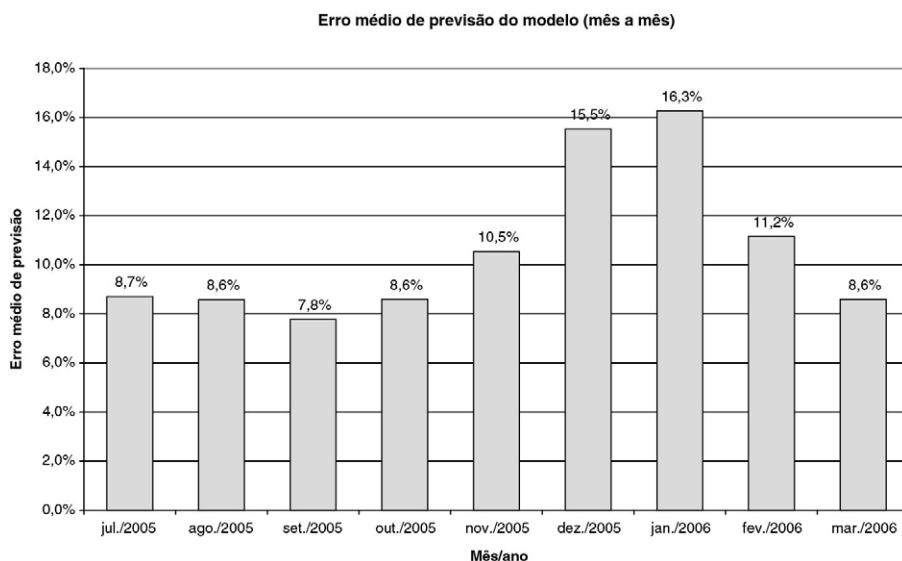


FIGURA 5.7 Erro médio mensal do modelo de regressão múltipla para prever a quantidade de ligações (julho de 2005 a março de 2006).

com os erros oriundos do módulo de previsão do software, com o intuito de verificar a ocorrência ou não de melhorias no processo de previsão.

Assim sendo, o MAPE do software para o mesmo período analisado foi de 13,54%, superior, portanto, aos 10,98% do modelo de regressão.

Além disso, uma análise mais detalhada dos erros de previsão do modelo proposto revela que alguns dos dias mais difíceis de prever e, portanto, com erros de previsão potencialmente maiores, consistem em feriados ou datas que se comportam quase como um feriado. Por exemplo, os dias 24/12/2005 e 31/12/2005 não foram classificados como feriados, mas certamente também não são dias comuns e devem gerar uma demanda de ligações inferior ao de um dia padrão. O problema é que o modelo de regressão não levou em conta isso, já que era possível classificar um dia apenas como sendo feriado ou não; e visto que esses dias não eram efetivamente feriados, foram tratados pelo modelo como dias comuns. Por este motivo, muito provavelmente, suas demandas foram superestimadas e seus erros de previsão foram bem acima da média: 34,2% e 37,2%, respectivamente, conforme já apresentado na [Tabela 5.6](#).

Problemas semelhantes, mas talvez em menor escala, certamente ocorreram em alguns dias do ano que consistiram em feriados locais, mas não nacionais. Como o modelo propõe uma previsão consolidada da demanda, só foi possível considerar os feriados nacionais; por exemplo, o dia 20/01/2006, um feriado local na Cidade do Rio de Janeiro, teve uma redução na demanda nacional provocada pela redução na demanda da cidade mas, ao ser tratado como um dia normal pelo modelo, teve sua demanda superestimada. Este problema também se observa nos feriados prolongados e nos dias que se situam entre um feriado e o fim de semana, em que a demanda foge do padrão.

Um problema ainda mais grave ocorreu com o dia 25/12/2005. Seu erro de previsão foi de quase 150%, já que a demanda prevista assumiu um valor negativo, -18.895 ligações, conforme a [Tabela 5.6](#). Isto aconteceu porque o dia em questão apresentava dois efeitos negativos e redutores de demanda, que foram somados pelo modelo de natureza aditiva: ser domingo e ser feriado. Mas, na prática, o efeito de um dia ser feriado certamente é reduzido se ele cair em um domingo ou sábado; e isto não foi levado em consideração, até porque só é possível considerar um dia como sendo feriado ou não, e não como uma espécie de feriado moderado.

Na verdade, esses problemas aumentaram os erros de previsão, mas isso não quer dizer que o processo futuro de previsão terá que incorrer nesses erros de grande magnitude quando se tratar de dias com comportamentos diferentes. Na prática, quando os analistas de previsão estiverem utilizando o modelo aqui proposto, eles poderão manipular os valores da variável *dummy* feriado para os dias diferentes do padrão.

Assim, por exemplo, o dia 24/12 pode receber manualmente o valor 0,7 para a sua variável Feriado, em vez de ser tratado como um dia normal; o dia 20/01 pode receber o valor 0,4 da mesma forma; um feriado que caia num domingo pode ter o seu valor da variável feriado reduzido de 1 para 0,6.

Depois dessa manipulação, basta considerar as informações do setor de faturamento da empresa telefônica sobre os ciclos de contas e usar a Equação 1 apresentada anteriormente para prever a quantidade de ligações para o produto em questão, a cada dia. Uma vez gerados os *outputs* da equação, os analistas podem incorporar os impactos

esperados pelos eventos especiais, como campanhas de mídias, novos serviços etc. e, desta forma, reduzir substancialmente os erros de previsão.

EXERCÍCIOS

1. A planilha **Exercício 1** apresenta algumas características de alguns imóveis que foram vendidos por uma corretora: a quantidade de banheiros, a área (em m^2) e o preço de venda (em US\$).
 - a. Desenvolva um modelo de Regressão Múltipla capaz de explicar o preço de venda do imóvel em função da quantidade de banheiros e da área.
 - b. Qual a capacidade explicativa deste modelo?
 - c. Como se comportam as variáveis independentes como explicadoras do preço de venda? Trabalhe com 10% de significância.
 - d. Aplique a equação gerada pelo modelo para “prever” o preço de venda de cada imóvel.
 - e. Calcule o erro percentual absoluto médio dessas previsões. Compare com o resultado encontrado na letra b.
 - f. Segundo o modelo, qual seria o preço de venda de um imóvel com dois banheiros e 200 m^2 ?
 - g. Segundo o modelo, qual seria o preço de venda de um imóvel com três banheiros e 200 m^2 ? Qual é o valor de um banheiro adicional?
 - h. Segundo o modelo, qual seria o preço de venda de um imóvel com três banheiros e 300 m^2 ? Qual é o valor de 100 m^2 adicionais?
2. Uma cadeia de lojas que vende CDs de música está tentando determinar a eficiência de sua propaganda via rádio e televisão. Foram coletados dados de algumas lojas da cadeia sobre vendas e gastos mensais em propaganda tanto por rádio quanto por televisão. Os dados estão na planilha **Exercício 2**.
 - a. Existe evidência, a 5% de significância, de alguma relação entre os gastos em cada tipo de propaganda e as vendas?
 - b. Que tipo de propaganda parece ser mais eficiente em relação ao custo?
 - c. Qual o nível de explicação do modelo como um todo?
3. Uma empresa que presta serviço de jardinagem está interessada em prever a quantidade mensal média que será gasta com serviços de jardinagem em uma certa vizinhança. Uma amostra de vizinhanças é selecionada, e a quantidade média gasta é registrada, juntamente com o tamanho médio dos terrenos (em lotes), a renda familiar média e o valor médio das casas. Os dados estão na planilha **Exercício 3**. Desenvolva e avalie um modelo de Regressão Múltipla para estimar o que impacta, a 5% de significância, os gastos médios por residência em serviços de jardinagem.

REFERÊNCIAS

ALAM, M. Using call centers to deliver public services. *House of Commons Paper*. London: The Stationery Office Books. 2002.

- BOUZADA, M. *O uso de ferramentas quantitativas em call centers – o caso Contax*. Orientador: Eduardo Saliby. Rio de Janeiro: UFRJ/COPPEAD, 2006. Tese (Doutorado em Administração de Empresas).
- CORRAR, L.; PAULO, E.; FILHO DIAS, J. *Análise multivariada para os cursos de Administração*. Ciências Contábeis e Economia. São Paulo: Atlas, 2007.
- GANS, N.; KOOLE, G.; MANDELBAUM, A. Telephone call centers: Tutorial, review and research prospects. *Manufacturing and Service Operations Management*, v. 5, 2, p. 79-141, 2003.
- GROSSMAN, T.; SAMUELSON, D.; OH, S.; ROHLER, T. Call centers *Encyclopedia of Operations Research and Management Science*. Boston: Kluwer Academic Publishers. 2001, p. 73-76.
- HALL, B.; ANTON, J. Optimizing your call center through simulation. *Call Center Solutions Magazine*, Oct 1998.
- HAWKINS, L.; MEIER, T.; NAINIS, W.; JAMES, H. Planning guidance document for US call centers. *Information Technology Support Center*, 2001 Maryland.
- KLUNGLE, R.; MALUCHNIK, J. The role of simulation in call center management. *MSUG Conference*, 1997.
- LEVINE, D.; BERENSON, M.; STEPHAN, D. *Estatística: Teoria e aplicações*. Rio de Janeiro : LTC.2000.
- MATAN, O.; NOURBAKHSH, I. Playing the numbers: Using ACD statistics for workforce management. *Telemarketing & Call Center Solutions*, Mar. 1998.
- MEHROTRA, V. Ringing up big business. *OR/MS Today*, 24, 4, p. 18-24, Aug. 1997.
- MEHROTRA, V.; PROFOZICH, D.; BAPAT, V. Simulation: the best way to design your call center. *Telemarketing & Call Center Solutions*, Nov. 1997.
- REYNOLDS, P. Forecasting fundamentals: The art and science of predicting call center workload. *Technology Marketing Corporation*, Feb 17, 2005 White Paper.
- SHEN, H.; HUANG, J. Analysis of call centre arrival data using singular value decomposition. *Applied Stochastic Models in Business and Industry*, v. 21, p. 251-63, 2005.
- SILVER, M. *Estatística para administração*. São Paulo: Atlas, 2000.
- WEINBERG, J.; BROWN, L.; STROUD, J. Bayesian forecasting of an inhomogeneous poisson process with applications to call center data. *University of Pennsylvania*, Jun 2006 White Paper.

Prevendo o tempo médio de atendimento em um *call center* por meio de um modelo de Regressão Múltipla

Marco Aurélio Carino Bouzada
MADE/UNESA

Objetivos de aprendizagem do capítulo

- Saber construir um modelo causal, capaz de explicar o comportamento de uma grandeza real com base no comportamento de outras variáveis.
- Entender as vantagens de um modelo causal sobre um modelo de série temporal.
- Aprender a quantificar variáveis qualitativas e a incluí-las no modelo por meio de variáveis *dummy*.
- Conseguir aferir a capacidade de explicação do modelo.
- Conseguir aferir a capacidade de explicação de cada variável e entender como cada uma impacta a variável dependente.
- Saber refinar o modelo, tornando-o mais parcimonioso com a retirada de variáveis pouco explicativas.
- Entender como obter e interpretar a equação capaz de prever a variável dependente.
- Aprender a mensurar o erro de previsão decorrente da utilização de modelo desenvolvido.

INTRODUÇÃO

Tendo sido conceituados os *call centers* no Capítulo 5, assim como informados os seus desafios operacionais, destacou-se que o custo de pessoal nessa indústria representa aproximadamente 70% do total. A matéria-prima para o processo de gerenciamento eficiente (níveis satisfatórios de serviço a custos razoáveis) de um *call center* consiste na previsão da demanda – tanto em termos da quantidade de ligações quanto do tempo de atendimento – e da consequente carga de trabalho.

Dessa forma, outra característica operacional que ganha importância e precisa ser estimada com acurácia para o correto dimensionamento da equipe de atendimento consiste no tempo médio de atendimento, também conhecida como TMA.

Uma previsão pode ser necessária para dimensionar as operações cotidianas ou para lidar com situações especiais, como, por exemplo, a implementação de uma nova tecnologia capaz de afetar não só o volume, mas também o padrão das chamadas.¹

Como o propósito real do processo de previsão é prever quantos agentes são necessários para lidar com o volume de chamadas, uma previsão precisa incluir – para ser completa – não apenas o volume de ligações, mas também a grandeza tempo de atendimento. Para prever o tamanho da equipe e os planos de horários posteriormente, é necessária uma fotografia da carga de trabalho, que consiste no número de chamadas multiplicado pelo TMA.²

O TMA é um dos alvos atacados pelo método proposto por Shen e Brown (2002), que é aplicado – entre outros propósitos – para modelar o padrão (dependente da hora do dia) do TMA das chamadas de clientes recebidas no *call center* de um banco israelense. O entendimento de tal comportamento variável desta grandeza é, na opinião dos autores, essencial para compreender o ambiente operacional de um sistema, e também para prever dinamicamente a sua carga de trabalho futura.

O PROCESSO DE PREVISÃO DO TMA NO CALL CENTER ESTUDADO SEM A REGRESSÃO MÚLTIPLA

Conforme dito no Capítulo 5, a Gerência de Planejamento de Tráfego da empresa preocupa-se sobretudo em conseguir melhorar as previsões de demanda, mas não só em termos da quantidade de ligações, como também em termos do TMA.

E a impressão da gerência quanto ao TMA é a mesma no que se refere ao volume de ligações: o tempo de atendimento para o produto 103 é influenciado pelos eventos de chegada da conta ao destino e pelo dia do seu vencimento, datas em que haveria um acesso diferenciado ao serviço por parte dos clientes.

O processo de previsão para o TMA segue os mesmos três passos, descritos no Capítulo 5, do processo de previsão do volume de chamadas. Mas os eventos especiais capazes de impactar a grandeza a ser prevista podem ser um pouco diferentes: mudanças no procedimento (roteiro de atendimento), que podem levar o operador a falar mais ou menos durante o atendimento; mudanças na Unidade de Resposta Audível (URA); entre outros. Mas, assim como no outro processo, aqui o comportamento dos ciclos das contas telefônicas dos clientes também consiste em uma informação que ajuda a previsão.

Sob os olhos de uma análise mais minuciosa, a magnitude do TMA tem variado muito ao longo do dia, assumindo, na parte da tarde, valores 35 a 40 segundos maiores que os tempos matutinos, segundo a coordenação da área. A impressão geral é que isso ocorre, principalmente, pela maior concentração de operadores novos na parte da manhã. Durante a noite, o TMA é ainda mais alto, principalmente – na opinião da equipe de previsão – por causa do menor nível de serviço oferecido neste período, o que pode fazer com que o cliente aguarde mais tempo na fila e reclame a respeito

1. Reynolds (2005).

2. Reynolds (2005); Wise (2006).

com o atendente antes de começar a tratar do assunto principal, propriamente dito, aumentando o tempo de atendimento.

O PROCESSO DE PREVISÃO DO TMA COM A REGRESSÃO MÚLTIPLA

A ideia novamente consiste em poder contar com uma ferramenta que use as informações sobre as contas dos clientes no processo de elaboração das previsões. Assim como anteriormente, o método escolhido foi o da Regressão Múltipla.

As informações em termos da data de chegada ao destino e data de vencimento das contas foram consolidadas e parcialmente já apresentadas nas Tabelas 5.1e 5.2, respectivamente.

Para desenvolver o modelo de regressão múltipla para prever o TMA em cada data, foi preciso ter acesso ao histórico desta grandeza. Por motivo de sigilo estratégico, esses valores não puderam ser revelados em sua plena magnitude; no entanto, a eles foi aplicada uma constante de proporcionalidade; os resultados, de julho de 2005 até março de 2006, fornecidos pela empresa, estão apresentados graficamente na [Figura 6.1](#).

Como se observa, a série não apresenta tendência de crescimento ou decréscimo, mas apenas um comportamento errático, dispensando, assim, a inclusão de uma variável associada à passagem do tempo.

De posse de todas essas informações foi possível montar uma planilha com os dados necessários para a construção do modelo de Regressão Múltipla dentro do qual o comportamento da variável dependente TMA em um dia específico poderia estar sendo explicado – inicialmente – por nove variáveis independentes: (i) dia da

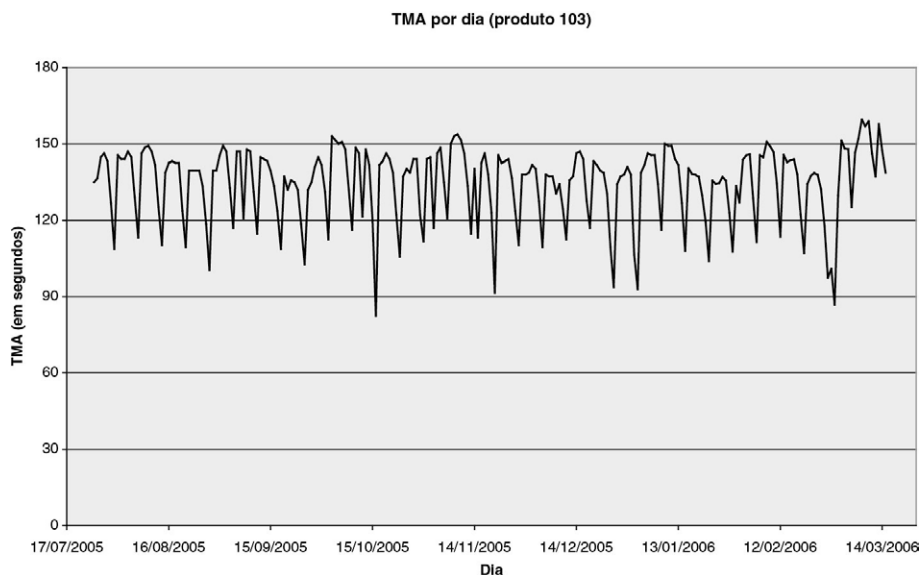


FIGURA 6.1 TMA por dia (julho de 2005 a março de 2006).

semana; (ii) feriado – se o dia específico consiste em um feriado ou não; (iii) chegada (C) – quantas contas estão previstas para chegarem ao cliente no dia específico; (iv) vencimento (V) – quantas contas vencem no dia específico; (v) $V - 3$ – para quantas contas aquele dia específico corresponde a três dias antes do vencimento, ou, em outras palavras, quantas contas irão vencer três dias depois da data específica; (vi) $V - 2$ – para quantas contas aquele dia específico corresponde à antevéspera do vencimento; (vii) $V - 1$ – para quantas contas aquele dia específico corresponde à véspera do vencimento; (viii) $V + 1$ – para quantas contas aquele dia específico corresponde a um dia depois do vencimento, ou, em outras palavras, quantas contas terão vencido na véspera daquela data específica; e (ix) $V + 2$ – para quantas contas aquele dia específico corresponde a dois dias depois do vencimento.

A Tabela 6.1 apresenta, em parte, a planilha de dados em questão.

Cabe salientar que as variáveis independentes dia da semana e feriado foram trabalhadas como sendo do tipo *dummy* (0 = Não; 1 = Sim). Foi criada uma variável *dummy* para a ocorrência de feriado e seis variáveis *dummy* para o dia da semana: domingo, segunda-feira, terça-feira, quarta-feira, quinta-feira e sexta-feira.

Dessa forma, para o dia 06/08/2005, por exemplo, a variável *dummy* feriado assumiu o valor 0; e as variáveis *dummy* domingo, segunda-feira, terça-feira, quarta-feira, quinta-feira e sexta-feira assumiram o valor 0, caracterizando, portanto e por exclusão, o dia em questão como sendo um sábado comum (não feriado).

Após a regressão ser executada, a primeira análise aponta para um modelo com um potencial de explicação bom, já que o R-quadrado ajustado indica que 79% da variação dos dados podem ser explicados. Neste caso, pode-se inferir a real influência das variáveis explicativas no tempo médio de atendimento e aceitar o modelo com elevada confiança, visto que o nível de significância da estatística F foi igual a 0,0000.

A contribuição de cada variável independente é expressa pelos coeficientes estimados, juntamente com os respectivos *p-values* (ou valores *p*), apresentados na Tabela 6.2. Observa-se um altíssimo valor *p* para a variável Chegada (C), indicando que a quantidade de contas chegando ao cliente naquele dia específico não pode ser vista como um bom previsor para o TMA na data em questão.

Apesar de ser uma variável crítica, ela deve ser retirada da análise, pois seu valor *p* está muito acima do razoável (10%) e porque o sinal do seu coeficiente angular (negativo) estaria indicando que o TMA em determinado dia diminuiria para cada conta que chegasse ao cliente no dia em questão, constituindo um efeito contrário ao esperado.

As variáveis $V - 3$ e $V - 2$ também apresentam valores *p* altos, indicando que a quantidade de contas para as quais aquele dia específico corresponde a dois ou três dias antes do vencimento não pode ser vista como um bom previsor para o TMA. O efeito pré-vencimento para o TMA parece, então, começar de forma significativa apenas na véspera ($V - 1$).

Assim como a variável Chegada (C), as variáveis $V - 3$ e $V - 2$ podem estar contaminando o modelo e convém retirá-las da análise. Sem ela, um novo modelo de regressão foi gerado, continuando com um bom potencial, já que 79% da variação dos dados podem ser explicados por ele.

TABELA 6.1 Planilha de dados do modelo de regressão múltipla para prever o TMA (agosto de 2005)

Dia	Dia da semana	Feriado?	Chegada (C)	V – 3	V – 2	V – 1	Vencimento (V)	V + 1	V + 2	TMA (seg)
01/08/2005	2ª-feira	Não	–	2142.215	–	–	2.136.197	–	–	145,50
02/08/2005	3ª-feira	Não	1.918.920	–	2.142.215	–	–	2.136.197	–	144,00
03/08/2005	4ª-feira	Não	2.177.552	–	–	2.142.215	–	–	2.136.197	144,00
04/08/2005	5ª-feira	Não	–	1.918.920	–	–	2.142.215	–	–	147,00
05/08/2005	6ª-feira	Não	218.438	–	1.918.920	–	–	2.142.215	–	144,75
06/08/2005	Sábado	Não	–	–	–	1.918.920	–	–	2.142.215	129,00
07/08/2005	Domingo	Não	–	–	–	–	1.918.920	–	–	113,25
08/08/2005	2ª-feira	Não	–	2.177.552	–	–	–	1.918.920	–	146,25
09/08/2005	3ª-feira	Não	2.036.848	–	2.177.552	–	–	–	1.918.920	148,50
10/08/2005	4ª-feira	Não	801.143	798.630	–	2.177.552	–	–	–	149,25
11/08/2005	5ª-feira	Não	–	–	798.630	–	2.177.552	–	–	147,00
12/08/2005	6ª-feira	Não	–	1.456.656	–	798.630	–	2.177.552	–	141,75
13/08/2005	Sábado	Não	–	–	1.456.656	–	798.630	–	2.177.552	125,25
14/08/2005	Domingo	Não	–	801.143	–	1.456.656	–	798.630	–	110,25
15/08/2005	2ª-feira	Não	–	–	801.143	–	1.456.656	–	798.630	138,75
16/08/2005	3ª-feira	Não	296.642	–	–	801.143	–	1.456.656	–	142,50
17/08/2005	4ª-feira	Não	246.901	296.642	–	–	801.143	–	1.456.656	143,25
18/08/2005	5ª-feira	Não	–	–	296.642	–	–	801.143	–	142,50

(Continua)

TABELA 6.1 Planilha de dados do modelo de regressão múltipla para prever o TMA (agosto de 2005) (*cont.*)

Dia	Dia da semana	Feriado?	Chegada (C)	V – 3	V – 2	V – 1	Vencimento (V)	V + 1	V + 2	TMA (seg)
19/08/2005	6ª-feira	Não	–	–	–	296.642	–	–	801.143	142,50
20/08/2005	Sábado	Não	–	246.901	–	–	296.642	–	–	123,75
21/08/2005	Domingo	Não	–	–	246.901	–	–	296.642	–	109,50
22/08/2005	2ª-feira	Não	–	–	–	246.901	–	–	296.642	139,50
23/08/2005	3ª-feira	Não	–	–	–	–	246.901	–	–	139,50
24/08/2005	4ª-feira	Não	2.136.197	–	–	–	–	246.901	–	139,50
25/08/2005	5ª-feira	Não	–	–	–	–	–	–	246.901	139,50
26/08/2005	6ª-feira	Não	–	–	–	–	–	–	–	133,50
27/08/2005	Sábado	Não	–	–	–	–	–	–	–	118,50
28/08/2005	Domingo	Não	–	–	–	–	–	–	–	100,50
29/08/2005	2ª-feira	Não	–	2.136.197	–	–	–	–	–	139,50
30/08/2005	3ª-feira	Não	2.181.734	–	2.136.197	–	–	–	–	139,50
31/08/2005	4ª-feira	Não	1.879401	–	–	2.136.197	–	–	–	145,50

TABELA 6.2 Coeficientes estimados das variáveis independentes do modelo de regressão múltipla para prever o TMA na data em questão

	Coeficientes	valor-P
Interseção	122,27	0,00%
Domingo	(15,17)	0,00%
2ª-feira	15,37	0,00%
3ª-feira	15,20	0,00%
4ª-feira	17,70	0,00%
5ª-feira	17,86	0,00%
6ª-feira	15,11	0,00%
Feriado	(29,20)	0,00%
Chegada (C)	(0,0000003)	68,35%
V – 3	0,0000008	29,97%
V – 2	0,0000011	16,89%
V – 1	0,0000016	3,44%
Vencimento (V)	0,0000029	0,01%
V + 1	0,0000026	0,12%
V + 2	0,0000022	0,36%

Os coeficientes de regressão estimados, juntamente com os respectivos valores p , estão apresentados na Tabela 6.3. Como pode ser observado, os valores p das outras variáveis não mudaram muito e, para este modelo modificado, apenas uma variável apresenta coeficiente angular com valor p superior (mas não muito) a 5%: V – 1. No entanto, será mantida no modelo, porque o valor p do seu coeficiente angular está abaixo de 10%.

Dessa forma, a retirada das variáveis C, V – 3 e V – 2 foi benéfica ao modelo, tornando-o mais parcimonioso e permitindo uma interpretação dos coeficientes obtidos.

O primeiro coeficiente corresponde ao valor da variável dependente TMA quando todas as variáveis independentes assumirem o valor zero; ou seja, quando domingo, segunda-feira, terça-feira, quarta-feira, quinta-feira, sexta-feira, feriado, V – 1, V, V + 1 e V + 2 forem zero –; quando o dia for sábado, não feriado e não consistir de nenhuma data dentro do intervalo de quatro dias em torno do vencimento de alguma conta, o TMA previsto é de 122,64 segundos.

Esse número pode ser considerado uma quantidade básica, à qual devem ser adicionados os efeitos da ocorrência das variáveis independentes, quantificados através dos seus respectivos coeficientes angulares.

TABELA 6.3 Coeficientes estimados das variáveis independentes do modelo (sem C; V – 3; V – 2) de regressão múltipla para prever o TMA na data em questão

	Coeficientes	valor-P
Interseção	122,64	0,00%
Domingo	(15,15)	0,00%
2ª-feira	15,36	0,00%
3ª-feira	14,96	0,00%
4ª-feira	17,33	0,00%
5ª-feira	17,84	0,00%
6ª-feira	15,03	0,00%
Feriado	(28,89)	0,00%
V – 1	0,0000012	6,96%
Vencimento (V)	0,0000031	0,00%
V + 1	0,0000032	0,00%
V + 2	0,0000024	0,04%

Assim, o fato de o dia ser domingo contribui com uma diminuição de 15,15 segundos no TMA previsto em relação ao número básico; o fato de o dia ser segunda-feira, contribui com um aumento esperado de 15,36 segundos no TMA; se for terça-feira, com um aumento de 14,96 segundos; se for quarta-feira, com um aumento de 17,33 segundos; se for quinta-feira, com 17,84 segundos a mais; e se for sexta-feira, contribui com um aumento de 15,03 segundos na previsão do tempo médio de atendimento, em relação ao dia básico.

Com essas informações, é possível elaborar um gráfico (Figura 6.2), que permite uma melhor visualização da sazonalidade do dia da semana. A evolução da demanda de ligações ao longo da semana, considerando dias sem a ocorrência de feriados e não próximos do vencimento de contas, começa com um valor baixo no domingo, apresentando um certo aumento na segunda-feira; a demanda vai se mantendo constante ao longo da semana, tendo uma queda razoável no sábado, antes de voltar ao patamar de domingo, ainda mais baixo.

De forma análoga aos dias da semana, o fato de um dia ser feriado reduz em 28,89 segundos o TMA previsto. Esta grande redução no tempo médio de atendimento tanto para feriados como para finais de semana quando comparados aos dias úteis é, de certa forma, estranha. Em princípio, não foi possível especular os motivos que levariam a tal variação. A equipe de previsão foi indagada a respeito e também não foi capaz de justificar o fato. A investigação acerca de tais motivos pode constituir-se em um interessante objeto de pesquisa futura.

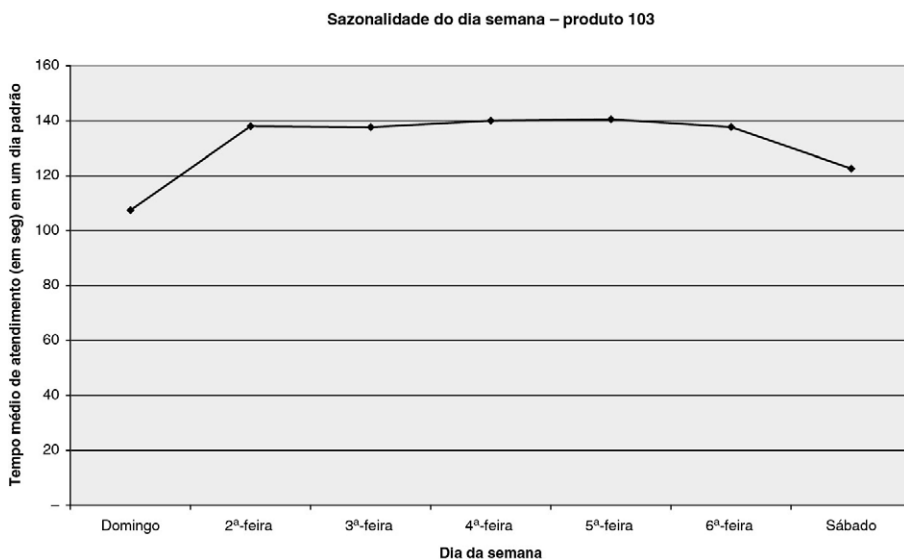


FIGURA 6.2 Comportamento sazonal do dia da semana – modelo sem C, V – 3 e V – 2.

Em relação aos dias que circundam a data do vencimento, os seguintes efeitos podem ser explicados, segundo o TMA: cada conta vencendo no dia seguinte (ou seja, para as quais a data em questão representa $V - 1$) aumenta em 0,0000012; cada conta vencendo no dia em questão aumenta em 0,0000031; cada conta vencendo na véspera aumenta em 0,0000032; e cada conta vencendo na antevéspera aumenta em 0,0000024.

Naturalmente, os coeficientes não têm apenas significado isolado, mas também podem e devem ser analisados em conjunto: se as consequências dos efeitos dia da semana, ocorrência de feriado e data crítica em relação ao vencimento da conta forem incorporadas ao valor previsto para um dia básico, será possível estabelecer uma equação para prever o TMA em determinado dia em função das características da data em questão. Dessa maneira, o TMA pode ser previsto conforme a [Equação 6.1](#).

As sete primeiras variáveis da equação são do tipo *dummy*, assumindo o valor 0 (não) ou 1 (sim). As quatro últimas representam, respectivamente: a quantidade de contas vencendo um dia depois; no próprio dia; um dia antes; e dois dias antes.

EQUAÇÃO 6.1 Equação do modelo para prever o TMA

$$\begin{aligned}
 \text{TMA (em segundos)} = & 122,64 - 15,15 \times \text{domingo} + \\
 & 15,36 \times \text{segunda-feira} + 14,96 \times \text{terça-feira} + \\
 & 17,33 \times \text{quarta-feira} + 17,84 \times \text{quinta-feira} + \\
 & 15,03 \times \text{sexta-feira} - 28,89 \times \text{feriado} + 0,0000012 \times V - 1 + \\
 & 0,0000031 \times V + 0,0000032 \times V + 1 + 0,0000024 \times V + 2
 \end{aligned}$$

De posse dessa equação, então, é possível calcular quanto o modelo de regressão teria previsto para o TMA em cada um dos dias acerca dos quais o histórico disponibiliza o valor real desta grandeza. Desta forma, pode-se comparar os valores que teriam sido previstos com os que efetivamente ocorreram, de maneira que um erro de previsão possa ser mensurado. Esta comparação pode ser visualizada na **Figura 6.3**.

A diferença entre os valores reais e os gerados pelo modelo também é conhecida como resíduo. Para ser usado um modelo de Regressão Múltipla, é necessário supor que a correlação entre os resíduos, ao longo do espectro das variáveis independentes, é zero; ou seja, que os resíduos são independentes entre si, não apresentando, consequentemente, autocorrelação serial.³

Uma forma de verificar a validade dessa suposição é através do teste de Durbin-Watson. No modelo que está sendo analisado, foi encontrado o valor 1,756 para a Estatística DW. Esta estatística obtida deve ser comparada com valores críticos que variam em função do nível de significância do teste, do tamanho da amostra e do número de variáveis independentes.⁴

Com 5% de significância, 234 dados na amostra e 11 variáveis independentes, o valor crítico inferior (d_L) obtido foi 1,688 e o valor crítico superior (d_U) encontrado foi 1,885.⁵

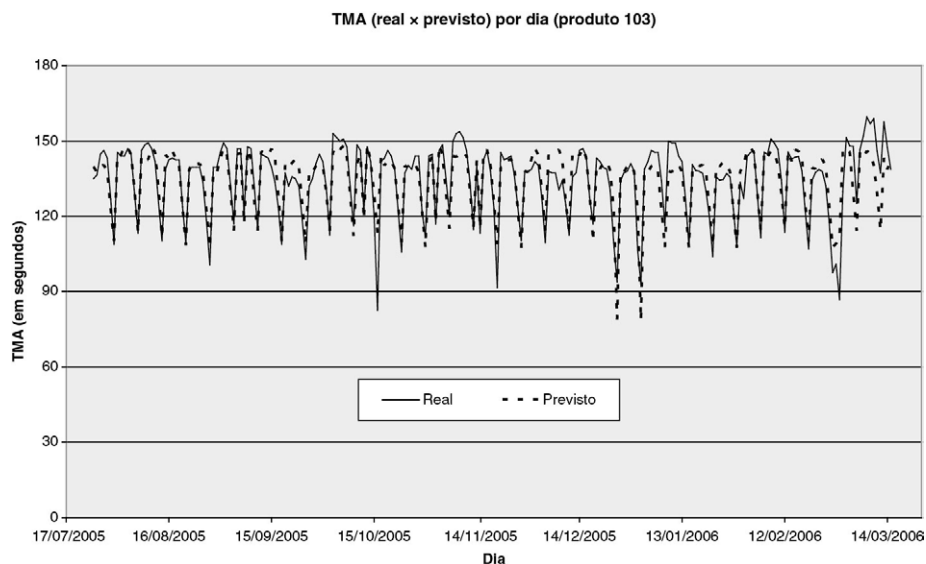


FIGURA 6.3 Valores reais multiplicados por uma constante e previstos pelo modelo para o TMA por dia (julho de 2005 a março de 2006).

3. Corrar; Paulo; Dias Filho (2007).

4. Corrar; Paulo; Dias Filho (2007).

5. Valores encontrados em tabela disponível no endereço <<http://www.stanford.edu/~clint/bench/dwcrit.htm>>. Acesso em 18/08/2009.



FIGURA 6.4 Resíduos do modelo X variável *dummy* Domingo.

Como o valor obtido para a Estatística DW (1,756) se situa entre os dois valores críticos encontrados, o teste para verificar a existência de autocorrelação dos resíduos é não conclusivo: não é possível garantir que os resíduos são independentes, mas também é incorreto afirmar que existe uma relação de dependência entre eles.⁶

Sendo assim, parece válido examinar o gráfico do resíduo contra cada variável independente a fim de procurar uma suposta autocorrelação e a identificação de uma eventual lei de formação nos valores dos resíduos à medida que o valor da variável independente avança. A Figura 6.4 ilustra o caso da variável *dummy* domingo. Como pode ser observado, não parece existir uma lei de formação; ou seja, a magnitude dos resíduos não parece depender do valor da variável independente (domingo).

Os resíduos apresentaram comportamento bem parecido em relação às outras variáveis *dummy* (feriado e as referentes aos outros dias da semana). E visto que o modelo trabalha com 11 variáveis independentes, seria, de certa maneira, inviável apresentar aqui tal quantidade de gráficos. Desta forma, será apresentada graficamente apenas a relação entre os resíduos e a outra variável independente mais marcante: Vencimento (Figura 6.5).

O comportamento dos resíduos em relação às outras variáveis ($V - 1$, $V + 1$ e $V = 2$) é bastante similar ao verificado na Figura 6.5: parecem estar distribuídos aleatoriamente em torno do eixo das abscissas, não apresentando nenhuma tendência.

Tudo isso gera uma segurança maior para os resultados do modelo, já que o teste para a verificação de autocorrelação serial dos resíduos foi não conclusivo. É claro que esta última análise é meramente visual, carecendo de um cientificismo maior. Mas é importante proceder dessa forma principalmente em situações em que os critérios

6. Corrar; Paulo; Dias Filho (2007).

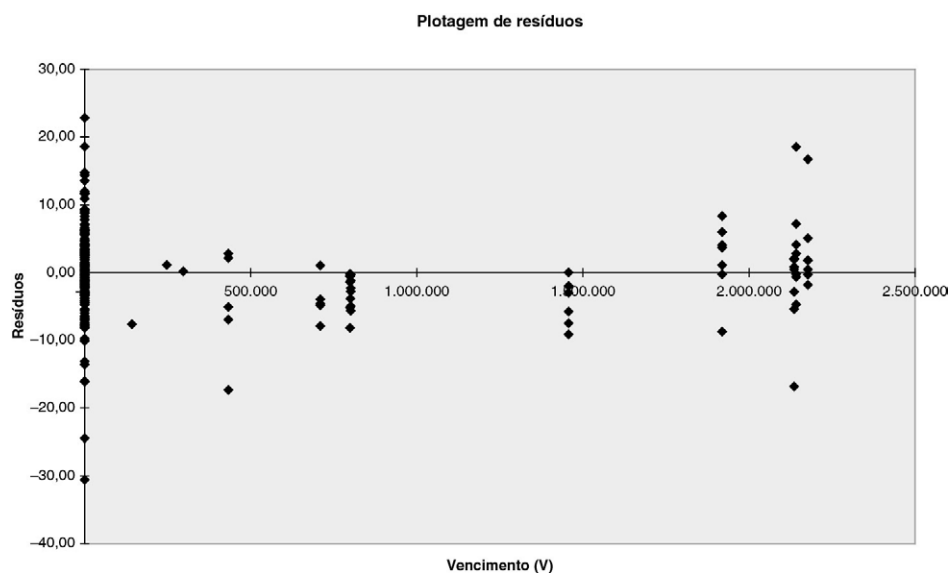


FIGURA 6.5 Resíduos do modelo X variável Vencimento.

objetivos apontam para uma indefinição, como foi o caso. Em relação à qualidade preditiva do modelo, foi possível perceber, na [Figura 6.3](#), excluindo alguns pontos específicos, uma boa aderência visual dos valores previstos em relação aos dados reais. Esta aderência foi aqui formalizada através do MAPE (Mean Absolute Percentual Error), ou erro médio percentual absoluto.

A título de ilustração, a [Tabela 6.4](#) apresenta dados relativos ao mês de janeiro de 2006, em que constam o TMA real, os valores que o modelo de regressão teria previsto para esta grandeza, assim como os erros percentuais absolutos oriundos desta previsão. Pode-se calcular que o modelo teria errado em 4,2%, em média, na previsão do TMA durante o mês em questão.

Mas janeiro é um mês atípico, com comportamento geralmente mais difícil de prever, sobretudo nos primeiros dias. Tal afirmação pode ser verificada neste exemplo particular a partir da [Figura 6.6](#), na qual é possível perceber que todos os outros meses relacionados, com exceção de dezembro, outro período atípico, apresentam erros médios de previsão menores do que o erro médio de janeiro de 2006.

Em média, o modelo está incorrendo em erro de previsão da ordem de 3,61%, valor satisfatoriamente baixo. Antes de compará-lo com o erro obtido pelo processo-padrão de previsão, é válido lembrar, por exemplo, que, durante este processo, os analistas recebem os valores previstos pelo software e acrescentam a eles os impactos esperados por eventos e premissas, como o comportamento do ciclo de contas e a ocorrência de feriados, assim como outros eventos especiais relacionados ao produto, como o plano de minutos, campanhas na mídia e outros potenciais aspectos impactantes.

Ora, o modelo de regressão proposto visa substituir apenas os valores sugeridos pelo software, acrescidos da influência dos feriados e do comportamento dos ciclos de contas; não se teve por pretensão substituir o papel do analista encarregado da previsão,

TABELA 6.4 Valores reais, previstos e erros do modelo de regressão múltipla para prever o TMA (janeiro de 2006)

Dia	Dia da semana	Feriado?	TMA (seg)	TMA (seg) Previsto pelo modelo)	Erro de previsão
01/01/2006	Domingo	Sim	93,00	78,60	15,5%
02/01/2006	2ª-feira	Não	138,75	137,99	0,5%
03/01/2006	3ª-feira	Não	141,75	137,60	2,9%
04/01/2006	4ª-feira	Não	146,25	139,97	4,3%
05/01/2006	5ª-feira	Não	145,50	140,48	3,5%
06/01/2006	6ª-feira	Não	145,50	137,67	5,4%
07/01/2006	Sábado	Não	134,25	122,64	8,6%
08/01/2006	Domingo	Não	116,25	107,49	7,5%
09/01/2006	2ª-feira	Não	150,00	137,99	8,0%
10/01/2006	3ª-feira	Não	149,25	137,60	7,8%
11/01/2006	4ª-feira	Não	149,25	139,97	6,2%
12/01/2006	5ª-feira	Não	144,00	140,48	2,4%
13/01/2006	6ª-feira	Não	141,75	137,67	2,9%
14/01/2006	Sábado	Não	126,75	122,64	3,2%
15/01/2006	Domingo	Não	108,00	107,49	0,5%
16/01/2006	2ª-feira	Não	140,53	137,99	1,8%
17/01/2006	3ª-feira	Não	138,33	137,60	0,5%
18/01/2006	4ª-feira	Não	137,94	139,97	1,5%
19/01/2006	5ª-feira	Não	137,13	140,48	2,4%
20/01/2006	6ª-feira	Não	129,92	137,67	6,0%
21/01/2006	Sábado	Não	119,52	122,64	2,6%
22/01/2006	Domingo	Não	103,92	108,36	4,3%
23/01/2006	2ª-feira	Não	135,62	140,21	3,4%
24/01/2006	3ª-feira	Não	134,42	139,84	4,0%
25/01/2006	4ª-feira	Não	134,69	141,70	5,2%
26/01/2006	5ª-feira	Não	137,07	140,48	2,5%
27/01/2006	6ª-feira	Não	135,58	137,67	1,5%
28/01/2006	Sábado	Não	122,99	122,64	0,3%
29/01/2006	Domingo	Não	107,68	107,49	0,2%
30/01/2006	2ª-feira	Não	133,53	137,99	3,3%
31/01/2006	3ª-feira	Não	127,10	140,22	10,3%

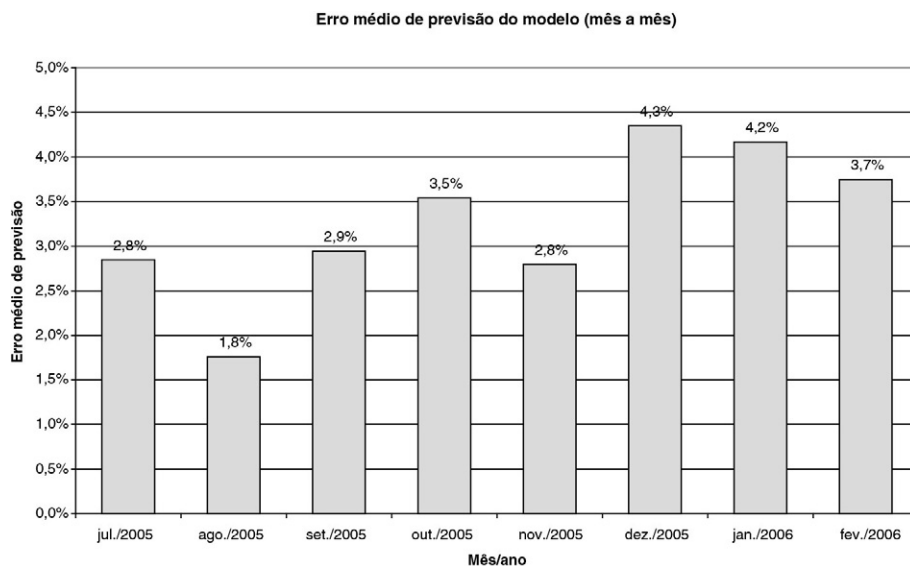


FIGURA 6.6 Erro médio mensal do modelo de regressão múltipla para prever o TMA (julho de 2005 a fevereiro de 2006).

dotado de toda a sua valiosa experiência e subjetividade, mas, sim, fornecer-lhe um subsídio complementar para a tomada de decisão. Assim, e já que o analista estará apto a acrescentar aos valores sugeridos pelo modelo de regressão os impactos que julgar relevantes, é mais justo comparar os erros de previsão do modelo aqui proposto com os erros oriundos do módulo de previsão do software, com o intuito de verificar a ocorrência ou não de melhorias no processo.

Assim sendo, o MAPE do software para o mesmo período analisado foi de 5,58%, razoavelmente superior, portanto, aos 3,61% do modelo de regressão.

Além disso, uma análise mais detalhada dos erros de previsão do modelo proposto revela que alguns dos dias mais difíceis de prever e, portanto, com erros de previsão potencialmente maiores, consistem em feriados ou datas que se comportam quase como um feriado. Por exemplo, os dias 24/12/2005 e 31/12/2005 não foram classificados como feriados, mas, certamente também não são dias comuns e devem gerar um TMA diferente daquele de um dia-padrão. O problema é que o modelo de regressão não levou em conta isso, já que era possível classificar um dia apenas como sendo feriado ou não; e já que esses dias não eram efetivamente feriados, foram tratados pelo modelo como dias comuns.

Problemas semelhantes, mas talvez em menor escala, certamente ocorreram em alguns dias do ano que consistiram em feriados locais, mas não nacionais. Como o modelo propõe uma previsão consolidada do TMA, só foi possível considerar os feriados nacionais; por exemplo, o dia 20/01/2006, um feriado local na Cidade do Rio de Janeiro, teve uma alteração no TMA nacional provocada pela alteração no TMA na cidade mas, ao ser tratado como um dia normal pelo modelo, teve seu TMA mal estimado. Este problema também se observa nos feriados prolongados

e nos dias que se situam entre um feriado e o fim de semana, em que o TMA foge do padrão.

Um problema ainda mais grave ocorreu com o dia 25/12/2005. Seu erro de previsão foi muito alto porque o dia em questão apresentava dois efeitos redutores de TMA, que foram somados pelo modelo de natureza aditiva: ser domingo e ser feriado. Mas, na prática, o efeito de um dia ser feriado certamente é reduzido se ele cair em um domingo ou sábado; e o modelo não levou isso em consideração, até porque só é possível considerar um dia como sendo feriado ou não, e não como uma espécie de feriado moderado.

Na verdade, esses problemas aumentaram os erros de previsão, mas isso não quer dizer que o processo futuro de previsão terá que incorrer nesses erros de grande magnitude quando ocorrerem dias com comportamentos diferentes. Na prática, quando os analistas de previsão estiverem utilizando o modelo aqui proposto, eles poderão manipular os valores da variável *dummy* feriado para os dias diferentes do padrão.

Assim, por exemplo, o dia 24/12 pode receber manualmente o valor 0,7 para a sua variável feriado, em vez de ser tratado como um dia normal; o dia 20/01 pode receber o valor 0,4 da mesma forma; um feriado que caia em um domingo pode ter o seu valor da variável feriado reduzido de 1 para 0,6.

Depois dessa manipulação, basta considerar as informações do setor de faturamento da Telemar sobre os ciclos de contas e usar a Equação 1 apresentada anteriormente para prever o TMA para o produto 103, a cada dia. Uma vez gerados os *outputs* da equação, os analistas podem incorporar os impactos esperados pelos eventos especiais, como implantação de novos serviços, mudanças no roteiro de atendimento etc. e, desta forma, reduzir substancialmente os erros de previsão.

EXERCÍCIOS

1. A planilha **Exercício 1** relaciona diversas características de 151 automóveis, assim como seus preços de mercado (em US\$).
 - a. Construa um modelo de Regressão Múltipla para explicar o comportamento da variável preço em função das variáveis relacionadas ao desempenho do carro.
 - b. O modelo ficou bom? Todas as variáveis são bons “explicadores lineares” do comportamento do preço, a 10% de significância?
 - c. Retire alguma(s) variável(is) que não esteja(m) explicando muito bem o preço e refaça a regressão. Avalie o novo modelo.
2. A companhia que vende leite deseja conhecer uma possível relação entre as quantidades vendidas e o valor investido em propaganda no mês. A planilha **Exercício 2** apresenta os valores relativos aos últimos 14 meses. Sabe-se que a empresa enfrentou uma greve do quinto ao sétimo mês.
 - a. Ignore a informação da greve e rode uma Regressão Simples para explicar a quantidade vendida em função do investimento em propaganda.
 - b. Refaça a Regressão Simples, eliminando os meses de greve. Avalie o resultado.
 - c. Trate a greve como uma variável *dummy* e faça uma Regressão Múltipla. Quanto seria vendido em um mês normal, em que fossem investidos R\$ 1.000 em propaganda?

3. A planilha **Exercício 3** mostra a demanda de energia e a temperatura (em °F) para diversos dias e períodos do dia em determinada região.
- Construa um modelo de Regressão Múltipla para prever a demanda horária de energia elétrica em função do período do dia (madrugada, manhã, tarde ou noite), do dia da semana e da Temperatura (°F). Avalie o modelo.
 - Quanto seria a demanda das 14 às 15 horas de uma quarta-feira, com temperatura média de 80°F?
 - E de 20 às 21 horas de um sábado, com temperatura média de 15°F?

REFERÊNCIAS

- CORRAR, L.; PAULO, E.; DIAS FILHO, J. *Análise multivariada para os cursos de Administração, Ciências Contábeis e Economia*. São Paulo: Atlas. 2007.
- REYNOLDS, P. Forecasting fundamentals: The art and science of predicting call center workload. Technology Marketing Corporation, White Paper, Feb. 17, 2005. <<http://www.tmcnet.com/channels/workforce-optimization/workforce-optimization-articles/forecasting-call-center-workload.htm>>.
- SHEN, H.; BROWN, L. Nonparametric regression and confidence bands with longnormal errors and application to bank call center data. In: JOINT STATISTICAL MEETINGS, 2002, New York. Proceedings. New York: JSM, 2002. p. 1-19.
- WISE, R. The Visible Difference: Making informed decisions with workforce management. IEX Corporation, White Paper, Nov. 2006. <<http://www.iex.com/white-papers/workforce-management-best-practices/the-visible-difference-making-informed-decisions-with-workforce-management.html>>.

Tutorial sobre Decomposição Clássica

Marco Aurélio Carino Bouzada

MADE/UNESA

Objetivos de aprendizagem do apêndice

- Aprender a prever os valores futuros de uma grandeza com base apenas no seu complexo comportamento histórico.
- Conhecer os componentes mais comuns das séries temporais.
- Saber isolar cada um desses componentes e entender a influência deles no comportamento da grandeza estudada.
- Aprender a estimar o comportamento futuro de cada um desses componentes.
- Conseguir agrupar todos esses componentes novamente para prever a grandeza estudada.
- Entender a influência do acaso no comportamento da série temporal e utilizar esse conhecimento para construir cenários pessimistas e otimistas de previsão.

Muitas grandezas apresentam um comportamento bastante irregular ao longo do tempo. Por exemplo, a venda de determinado produto ao longo de um mês em determinada filial de uma loja; ou a temperatura ao longo do dia em um determinado bairro; ou a intensidade pluviométrica ao longo do ano em uma determinada cidade; entre outras.

Ao observarmos esse comportamento irregular (através de gráficos, por exemplo), podemos imaginar a dificuldade em encontrar um padrão de comportamento para a grandeza em voga. Sem identificar esse padrão de comportamento, fica muito difícil, por exemplo, fazer um exercício de previsão acerca do futuro da grandeza.

Imaginemos que o **Gráfico 1** representa o consumo de energia elétrica no Brasil ao longo das semanas. Esse comportamento irregular pode parecer desprovido de padrão mas, como veremos mais adiante, não é exatamente o que ocorre.

O método apresentado a partir de agora é conhecido como **Decomposição Clássica** e tem como objetivo entender o comportamento (decompondo-o) das grandezas ao longo do tempo, de modo a auxiliar na previsão para os valores futuros das grandezas.

A irregularidade no comportamento das grandezas normalmente é resultado de diversos componentes. No exemplo do consumo de energia elétrica (**Gráfico 1**), esta

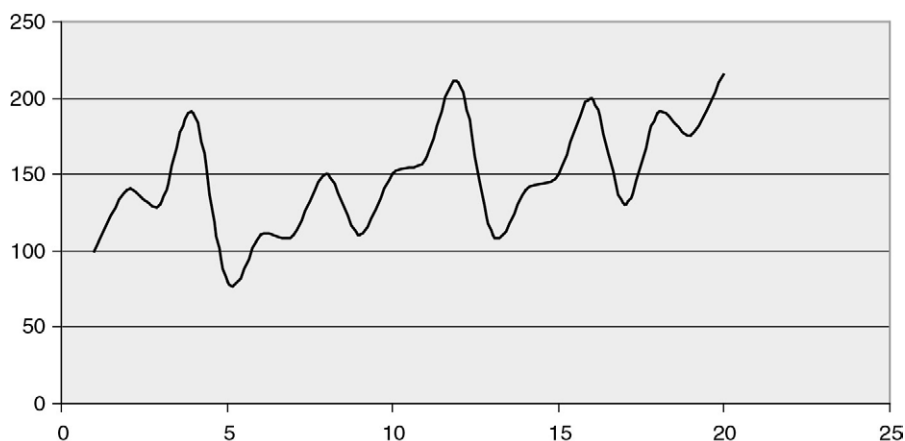


GRÁFICO 1 Consumo de energia elétrica no Brasil.

grandeza varia ao longo do tempo em função da sazonalidade, da tendência, do ciclo e de componentes aleatórios. Estes componentes aparecem em parte dos comportamentos irregulares apresentados pelas grandezas. Nesses casos, a Decomposição Clássica é o método de previsão mais adequado, já que é nesses elementos que as grandezas são decompostas neste método. A seguir, esses componentes são explicados.

SAZONALIDADE (S)

Explica o comportamento cíclico (ao longo de um ano ou intervalo menor) das grandezas, muitas vezes sendo resultado de suas variações em função das estações do ano (por exemplo: venda de sorvetes, consumo de energia elétrica). Nestes casos, o padrão de sazonalidade é mensal ou trimestral ao longo do ano. A venda de sorvetes no Brasil, por exemplo, é bem grande no primeiro trimestre de cada ano, diminui no segundo, atinge seu menor valor no terceiro e volta a aumentar no quarto.

Mas a sazonalidade pode se manifestar de forma diferente da mensal/semanal e/ou não ser fruto das variações climáticas. Exemplos:

- as vendas semanais de uma empresa ao longo do mês – que podem ser maiores nas últimas semanas e menores nas primeiras em razão do cumprimento de cotas por partes dos vendedores;
- consumo diário de bebida alcoólica nos bares ao longo da semana – que é maior nos últimos dias da semana e menor nos primeiros por causa do movimento nos bares;
- a intensidade de tráfego por hora ao longo do dia – que apresenta picos nos horários de “rush” – de manhã cedo e no fim da tarde.

TENDÊNCIA (T)

As grandezas, em boa parte das vezes, apresentam uma tendência de comportamento em relação à variável tempo. Assim, nesses casos, elas revelam um comportamento que tende a ser maior (na maioria das vezes) ou menor à medida que o tempo passa. Esta

tendência, normalmente, é consequência do aumento da população e/ou da evolução tecnológica/científica, mas também pode ser fruto de outras causas, como o aumento gradativo da temperatura na Terra ou até mesmo de mudanças no comportamento cultural das sociedades.

Dessa forma:

- a receita de vendas de uma empresa tende a crescer por causa do aumento da população;
- o percentual de óbitos em uma cidade tende a diminuir por causa da evolução científica;
- o percentual de mulheres trabalhando fora de casa tende a aumentar por causa de modificações culturais.

E, por causa disso, é de se esperar, *a priori*, que a receita de vendas de uma empresa tenha sido maior em 2001 do que em 1982; e que o percentual de óbitos em uma cidade tenha sido menor em 2010 do que em 1950.

CICLO (C)

Parte do comportamento irregular das grandezas pode ser atribuída ao ciclo, que se manifesta na forma de flutuações semelhantes a ondas em torno das tendências, que têm duração maior do que um ano, mas que não têm a exata periodicidade que caracteriza o padrão sazonal. A identificação das épocas em que ocorrem os picos e depressões destas ondas sugere seu relacionamento com a flutuação geral da atividade econômica – o chamado ciclo econômico.

Portanto, mesmo que uma grandeza não revele sazonalidade nem tendência quando observada a sua variação em relação ao tempo, ainda assim pode apresentar um comportamento cíclico não aleatório, que pode ser visualizado, entendido e decomposto de modo a ser usado num futuro exercício de previsão. O montante *per capita* de dinheiro gasto em restaurantes em São Paulo, por exemplo, não apresenta, a princípio, nenhum padrão sazonal e nenhuma tendência de crescimento ou redução ao longo do tempo; mas seu comportamento pode ser explicado pela variação do nível econômico e ter sua variação atribuída exclusivamente ao ciclo (além de a componentes aleatórios). Se for esperada uma atividade econômica de intensidade em 2006 maior do que em 1999 e menor do que em 2003, há de se prever que a grandeza em questão seja maior em 2003 e menor em 1999 (com um valor intermediário para 2006), independente do mês do ano.

COMPONENTES ALEATÓRIOS (A)

Se expurgarmos os efeitos sazonais, cíclicos e de tendência de uma série que revele o comportamento de uma grandeza ao longo do tempo, o restante da sua variação é denominado flutuação residual e pode ser considerado como proveniente de componentes aleatórios. Esta flutuação residual também pode ser entendida como oriunda de situações indistinguíveis, já que a sazonalidade, a tendência e o ciclo foram considerados como situações distinguíveis.

Como dito anteriormente (e será detalhadamente explicado a partir de agora, com o auxílio do exemplo da venda de lubrificantes no Brasil e da planilha **Decomposição**

clássica.xls), o método de Decomposição Clássica isola os efeitos (sazonalidade, tendência, ciclo e componentes aleatórios) que geram a irregularidade do comportamento das séries temporais, prevê seu comportamento (de cada efeito, isoladamente) no futuro e, novamente, agrupa esses efeitos de forma a obter a melhor previsão para o futuro da série temporal.

A série temporal pode, então, ser representada da seguinte forma:

$$R = S \times T \times C \times A$$

em que **R** é o valor real da grandeza e os fatores são os efeitos mencionados anteriormente

O MÉTODO DA DECOMPOSIÇÃO CLÁSSICA

Ao abrirmos a planilha **Decomposição clássica.xls**, podemos observar a série de dados históricos que representa a venda de lubrificantes no Brasil, por trimestre, entre 1983 e 1999. O primeiro passo para a elaboração de uma previsão baseada em dados passados é **fazer um gráfico da série histórica**, de modo a melhor perceber o seu comportamento. Fazemos isso pressionando o botão **Executar passo da planilha**.

Ao observarmos o gráfico, podemos verificar uma flutuação muito grande dos dados, além de uma tendência relativamente nítida de crescimento. Para decompor a série nos seus quatro efeitos já mencionados – tendência, ciclo, componentes aleatórios e sazonalidade –, vamos começar por este último.

Os dados são apresentados em trimestres e parece que é nessa unidade de tempo que a sazonalidade se manifesta. Se compararmos as vendas totais do ano de 1984 com as vendas totais dos anos de 1985, podemos entender que essa comparação está livre de sazonalidade, porque ambos os anos têm um primeiro trimestre, um segundo trimestre, um terceiro trimestre e um quarto trimestre, eliminando as diferenças devidas a esse efeito. Na verdade, considerando o intervalo de um ano, boa parte do efeito aleatório é anulada, porque os componentes positivos e negativos tendem a se anular ao longo de um ano.

Entretanto, é interessante continuarmos trabalhando na mesma grandeza dos dados originais – a venda trimestral. Para tal, não lidaremos com a venda anual, mas, sim, com a média móvel de quatro períodos (trimestres). Assim, a primeira média móvel será a média de vendas do primeiro ao quarto trimestres; a segunda média móvel será a média de vendas do segundo ao quinto trimestres; e assim por diante. Podemos observar que cada uma dessas médias móveis contém um primeiro trimestre, um segundo trimestre, um terceiro trimestre e um quarto trimestre, anulando o efeito da sazonalidade.

Nosso próximo passo é, então, **calcular essas médias móveis**, pressionando novamente o botão **Executar passo** da planilha (esse procedimento deverá ser repetido sempre que estiver sendo indicado o próximo passo). A fórmula das médias móveis pode ser observada na planilha. A rigor, a primeira média móvel se situa entre o segundo e o terceiro trimestres (ponto médio do intervalo), a segunda entre o terceiro e o quarto trimestres, e assim por diante. Para efeito de apresentação, elas foram colocadas “meia linha” acima na planilha, ou seja, a primeira média móvel na segunda linha, a segunda média móvel na terceira linha, e assim por diante.

Para fazer esse ajuste, é necessário avançar para o passo seguinte, que consiste em **calcular as médias móveis centradas**. A média móvel centrada corresponde à média entre duas médias móveis consecutivas e se situa no ponto médio entre elas (sua fórmula pode ser observada). Assim, a primeira média móvel centrada se situa na terceira linha (já que consiste na média das duas primeiras médias móveis – situadas entre a segunda e a terceira linhas, e entre a terceira e a quarta linhas, respectivamente); a segunda média móvel centrada na quarta linha; e assim por diante.

O próximo passo é **plotar as médias móveis centradas no gráfico**, de modo a observar como seu comportamento está livre da sazonalidade (e de boa parte dos componentes aleatórios também). Observamos que a flutuação da série diminuiu bastante e podemos concluir que a sua variação é fruto da tendência e do ciclo econômico. Esta conclusão pode ser mais bem entendida – ao considerarmos que a média móvel consiste nos dados originais (R), sem o efeito da sazonalidade e dos componentes aleatórios – através da fórmula:

$$\text{Média Móvel} = \frac{S \times T \times C \times A}{S \times A} = T \times C$$

Como se trata de um exercício de decomposição, está na hora de decompor a média móvel ($T \times C$) em tendência (T) e ciclo (C). Para tal, devemos executar o próximo passo, que é **determinar a tendência**. Isso é feito a partir da opção (do Excel) de adicionar uma linha de tendência à série da média móvel centrada. A reta plotada é crescente (revelando uma tendência positiva, provocada, provavelmente, pelo aumento da população brasileira e do seu poder aquisitivo no período da série) e pode ser representada pela equação $y = 61,324x - 118.654$. Esta reta explica, aproximadamente, 72% dos dados ($R^2 = 0,7238$; que pode ser considerado um bom nível de explicação) que compõem a série das médias móveis centradas.

O próximo passo é **quantificar os valores dessa reta de tendência**, de modo que possamos ter dados para cada um dos trimestres do período considerado. Essa quantificação é realizada por meio da equação da reta, anteriormente apresentada, que pode ser observada nas fórmulas da planilha.

Essa quantificação é importante como preparação para o próximo passo, a **determinação dos resíduos cíclicos**. Como sabemos, a média móvel consiste nos efeitos da tendência e do ciclo. Portanto, se eliminarmos o efeito da tendência (T), teremos apenas os resíduos cíclicos (C):

$$\frac{\text{Média Móvel}}{T} = \frac{T \times C}{T} = C$$

Essa relação também pode ser observada nas fórmulas da planilha. Esses resíduos cíclicos consistem, na verdade, de flutuações em torno da tendência, resultantes da maior ou menor intensidade do ciclo econômico. Assim, o primeiro trimestre de 1986 viveu uma situação de baixa intensidade do ciclo econômico (88,7%), reduzindo em 11%, aproximadamente, o nível de vendas projetado pela tendência para esse período. Analogamente, o segundo trimestre de 1989 experimentou uma forte intensidade econômica (109,3%), aumentando em 9%, aproximadamente, o nível de vendas projetado pela tendência para este período.

Essas flutuações podem ser mais bem evidenciadas depois de executado o passo seguinte: **elaboração do gráfico dos resíduos cíclicos**. Nesse gráfico, a tendência está representada pela linha horizontal de 100% e os resíduos podem ser observados como flutuações em torno dessa linha. Assim, pode-se ver que o pico econômico ocorreu no terceiro trimestre de 1965 (118,7%) e o menor nível do ciclo aconteceu no terceiro trimestre de 1983 (77,8%).

Com a média móvel, reunimos os efeitos da tendência (T) e do ciclo econômico (C), conforme já constatado. Se dividirmos os valores originais (R) pelos valores da média móvel, ficaremos apenas com o efeito da sazonalidade e dos componentes aleatórios, já que:

$$\frac{R}{\text{Média Móvel}} = \frac{S \times T \times C \times A}{T \times C} = S \times A$$

E é esse o próximo passo: **a avaliação do efeito da sazonalidade e dos componentes aleatórios**. As fórmulas podem ser observadas na planilha.

Assim como fizemos anteriormente, ao separarmos os efeitos da tendência dos efeitos do ciclo, nós o faremos com o efeito da sazonalidade e dos componentes aleatórios. Começaremos calculando os índices sazonais (são quatro: um para cada trimestre), que representam variações relativas entre os trimestres. A média de todos os índices sazonais deve ter sempre o valor 1, já que, ao considerarmos o período de sazonalidade (neste caso, o ano inteiro), o efeito sazonal deve sumir (conforme já mencionado).

Ao agruparmos todos os valores (de $S \times A$) por trimestre e tirarmos sua média, obteremos os índices sazonais, pois os componentes aleatórios tendem a se anular ao juntarmos 15, 20 valores, e o resultado é proveniente apenas de efeitos sazonais. Na verdade, é conveniente, no cálculo dessas médias, eliminar os valores extremos (o menor e o maior) dos grupos de modo a se evitar considerar *outliers* (originados por fatores como greves e incêndios), já que um desvio muito grande pode distorcer uma média adequada.

Assim, o índice sazonal do primeiro trimestre consiste na média dos valores (de $S \times A$) dos primeiros trimestres (com exceção do menor e do maior). A fórmula é análoga para os outros trimestres e pode ser observada na planilha após a execução do passo seguinte (**cálculo dos índices sazonais**). Esses índices se repetem, já que existe apenas um índice sazonal para o primeiro trimestre, um índice sazonal para o segundo trimestre, e assim por diante.

Como a média utilizada é modificada (em função da exclusão dos valores extremos), é necessário executar o próximo passo e **ajustar os índices sazonais**. Este ajuste consiste em dividir cada índice sazonal pela média dos quatro índices (conforme fórmulas na planilha) de modo a fazer com que sua média seja igual ao valor 1. Pelo fato de as médias não terem contido os valores extremos, a média dos índices sazonais ficou, nesse caso, um pouco maior do que 1 (1,001), devendo os índices ser ajustados (reduzidos) por esse fator.

Nesse momento, podemos observar que os maiores índices sazonais ocorrem no segundo e no quarto trimestres, respectivamente. Esse comportamento pode ser entendido através do fato de as distribuidoras de petróleo trabalharem com cotas semestrais de vendas, deslocando os picos de venda para os finais dos semestres.

Analogamente à separação dos efeitos do ciclo dos efeitos da tendência, a separação da sazonalidade dos componentes aleatória é mais bem visualizada graficamente. O passo seguinte consiste em **elaborar os gráficos dos componentes aleatórios**, um para cada trimestre. Assim, no gráfico do primeiro trimestre, a reta horizontal representa o índice sazonal do primeiro trimestre e a flutuação dos valores (de $S \times A$) em torno dela é explicada exclusivamente por componentes aleatórios. Para os outros trimestres, a ideia é a mesma.

Naturalmente, o próximo passo é **calcular os componentes aleatórios**, na forma de uma divisão dos valores (de $S \times A$) pelos índices sazonais (S), conforme as fórmulas na planilha e a seguir:

$$\frac{R}{\text{Média Móvel}} \times \frac{1}{S} = \frac{S \times A}{S} = A$$

Esses componentes aleatórios devem ser mais bem observados em uma **distribuição acumulada de probabilidade** (cuja elaboração consiste no passo seguinte), já que serão utilizados na confecção de uma previsão probabilística de vendas, conforme veremos mais adiante. Assim, ao analisarmos alguns pontos notáveis desse gráfico, podemos concluir que: existem 47% de probabilidade de o componente aleatório ser menor do que 100%, ou seja, de o valor real futuro de venda ser inferior ao da previsão pontual¹ (desconsiderando os componentes aleatórios); e que existem 5% de probabilidade de o componente aleatório ser maior do que 109%.

Para fazermos uma previsão de vendas para o futuro, imaginaremos que o padrão sazonal se manterá no período previsto. Assim, o próximo passo é **repetir os índices sazonais para os trimestres futuros**.

É de se esperar, também, que a tendência se manifeste – no período previsto – da mesma forma que se manifestou no passado. Desta forma, o passo seguinte consiste na **extrapolação da tendência futura**. Podemos observar que os valores da tendência são resultado da mesma fórmula que compôs os valores passados.

O próximo passo (**preenchimento dos valores futuros do ciclo econômico**) requer um pouco de informações, bom senso e imaginação por parte do previsor, já que não há motivos para crer que essa variável apresente, no futuro, o mesmo padrão de comportamento do passado. Desta forma, o previsor pode consultar especialistas da área econômica, ou adotar um valor único que represente a intensidade com a qual o ciclo econômico se manifestará no período previsto (este valor pode ser a média dos valores históricos ou um outro valor que o previsor julgue ser mais representativo do comportamento futuro). No nosso exemplo, o previsor reuniu informações que o levaram a considerar que:

- o ano de 2000 apresentaria um nível econômico um pouco mais intenso (88%) que o de 1999;
- esse nível cairia para 83% em 2001;
- esse nível subiria para 95% em 2002.

1 A previsão de vendas pontual ignora os componentes aleatórios, é realizada compondo-se os efeitos de sazonalidade, tendência e ciclo e será elaborada a partir de agora, para o período de 2000 a 2002 (três anos ou 12 trimestres).

A **composição das vendas futuras** é executada no passo seguinte, simplesmente agrupando (compondo) os três fatores: $T \times C \times S$, fórmula que pode ser observada na planilha. Essa é a previsão pontual, que ignora o efeito dos componentes aleatórios.

A previsão probabilística de vendas considera esse efeito (e reproduz da mesma forma que na pontual o comportamento dos outros três efeitos) e será ilustrada por três cenários:

- **Pessimista**, confeccionado no passo seguinte, em cor vermelha, para o qual o componente aleatório escolhido foi o segundo menor (para retirar o valor extremo) entre os dados históricos (84,9%). Ao observarmos esse valor na distribuição acumulada de probabilidade, podemos perceber que vendas reais inferiores a esses valores previstos têm menos de 3% de probabilidade de ocorrer.
- **Otimista**, confeccionado no próximo passo, em cor verde, para o qual o componente aleatório escolhido foi o segundo maior (para retirar o valor extremo) entre os dados históricos (110,3%). Ao observarmos esse valor na distribuição acumulada de probabilidade, podemos perceber que vendas reais superiores a esses valores previstos têm, aproximadamente, 5% de probabilidade de ocorrer.
- **Meio-termo**, confeccionado no passo seguinte, em cor amarela, para o qual o componente aleatório escolhido foi o valor 104,2%. Ao observarmos esse valor na distribuição acumulada de probabilidade, podemos perceber que vendas reais inferiores a esses valores previstos têm, aproximadamente, 72% de probabilidade de ocorrer.

O último passo consiste em **plotar esses cenários (juntamente com a previsão pontual) no gráfico de vendas**, de modo a comparar os possíveis comportamentos futuros dessa grandeza com o seu comportamento histórico. O previsor pode e deve, inclusive, observar entre que faixas de valores as vendas se situarão no futuro, e com que probabilidade de ocorrência.

EXERCÍCIOS

1. A planilha **Exercício 1** relaciona as vendas mensais de todo o varejo dos Estados Unidos (em bilhões de dólares) de janeiro de 1984 a julho de 1993. Faça a previsão do comportamento dessa grandeza para os dois anos seguintes.
2. A planilha **Exercício 2** relaciona as vendas mensais de uma cadeia de lojas de departamento de janeiro de 1993 a dezembro de 1995 no Brasil. Faça a previsão do comportamento dessa grandeza para os dois anos seguintes.
3. A planilha **Exercício 3** relaciona as vendas mensais (em unidades) de uma cadeia de lojas de artigos femininos de janeiro de 1986 a dezembro de 1992 no Brasil. Faça a previsão do comportamento dessa grandeza para os dois anos seguintes.
4. A planilha **Exercício 4** relaciona as vendas mensais (em 1.000m³) de óleo diesel no Brasil de janeiro de 1984 a setembro de 1994. Faça a previsão do comportamento dessa grandeza para os dois anos seguintes.

5. A planilha **Exercício 5** relaciona a produção mensal (em 1.000 litros) de cerveja na Austrália de janeiro de 1956 a agosto de 1995. Faça a previsão do comportamento dessa grandeza para os dois anos seguintes.

REFERÊNCIAS

- BOUZADA, M. Ensinando a decomposição clássica de forma não tão clássica. In: ENCONTRO DE ENSINO E PESQUISA EM ADMINISTRAÇÃO E CONTABILIDADE (ENEPQ), 2., 2009, Curitiba. Anais... Curitiba: ANPAD, 2009.
- TAVARES, L.; OLIVEIRA, R.; THEMIDO, I.; CORREIA, F. *Investigação operacional*. Lisboa: McGraw-Hill. 1996.